

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 16:17:19 ; Search time 71.77 Seconds
(without alignments)
253.563 Million cell updates/sec

Title: US-09-052-089a-2

Perfect score: 2393
Sequence: 1 MFILSLCTICSDFFDHSRDV.....VRKTVSSASQPKLDTFLCQ 470

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 198 | 8.3 | 1976 | 1 MYHA_BOVIN | Q27991 Bos taurus |
| 2 | 192.5 | 8.0 | 1325 | 1 G160_MOUSE | P59337 mus musculus |
| 3 | 190 | 7.9 | 2017 | 1 MYSN_DROME | Q99323 drosophila |
| 4 | 185.5 | 7.8 | 1102 | 1 MYSC_CHICK | P29616 gallus gall |
| 5 | 185.5 | 7.8 | 1940 | 1 MYH3_RAT | P12847 rattus norv |
| 6 | 185.5 | 7.8 | 1976 | 1 MYH3_RAT | Q9J110 rattus norv |
| 7 | 185.5 | 7.8 | 2116 | 1 MYH2_DICDI | P08799 dictyosteli |
| 8 | 184.5 | 7.7 | 1940 | 1 MYH3_HUMAN | P11055 homo sapien |
| 9 | 183 | 7.6 | 1938 | 1 MYH3_AEOR | P24733 aequipecten |
| 10 | 182.5 | 7.6 | 501 | 1 MYSD_RABIT | Q99105 oryctolagus |
| 11 | 180 | 7.5 | 1938 | 1 MYHD_HUMAN | Q9UKX2 homo sapien |
| 12 | 180 | 7.5 | 1969 | 1 MYSA_CAEEL | P12844 caenorhabdi |
| 13 | 179.5 | 7.5 | 1976 | 1 MYH4_HUMAN | P25663 caenorhabdi |
| 14 | 178 | 7.4 | 1790 | 1 USOL_YEAST | P35366 saccharomyc |
| 15 | 177.5 | 7.4 | 1957 | 1 YD86_SCHPO | Q10411 schizosacch |
| 16 | 177 | 7.4 | 1960 | 1 MYH9_HUMAN | P35579 homo sapien |
| 17 | 176.5 | 7.4 | 895 | 1 RASO_TREVO | P58302 thermoplas |
| 18 | 176.5 | 7.4 | 1509 | 1 MTSN_ACACA | P05656 acanthamoeb |
| 19 | 176 | 7.4 | 1084 | 1 MYSS_RABIT | P02563 oryctolagus |
| 20 | 175 | 7.3 | 882 | 1 MYSP_CAEEL | P10567 caenorhabdi |
| 21 | 174.5 | 7.3 | 1961 | 1 MYH9_RAT | Q62812 rattus norv |
| 22 | 173.5 | 7.3 | 1203 | 1 XCOPE_XENLA | P30533 xenopus lae |
| 23 | 172.5 | 7.2 | 1937 | 1 MYH8_HUMAN | P13535 homo sapien |
| 24 | 171.5 | 7.2 | 3210 | 1 CENF_HUMAN | P19454 homo sapien |
| 25 | 171 | 7.1 | 1130 | 1 YL17_CAEEL | Q11102 caenorhabdi |
| 26 | 170.5 | 7.1 | 1939 | 1 MYH4_HUMAN | Q9Y623 homo sapien |
| 27 | 170.5 | 7.1 | 2871 | 1 DESP_HUMAN | P15924 homo sapien |
| 28 | 170 | 7.1 | 1093 | 1 TMFL_HUMAN | P82094 homo sapien |
| 29 | 169 | 7.1 | 995 | 1 H1P1_HUMAN | O00281 homo sapien |
| 30 | 169 | 7.1 | 1238 | 1 Y009_CAEEL | O09236 caenorhabdi |
| 31 | 169 | 7.1 | 1935 | 1 MYSS_CYPCA | Q90339 cypripus ca |
| 32 | 169 | 7.1 | 1959 | 1 MYH9_CHICK | P14105 gallus gall |
| 33 | 168.5 | 7.0 | 1012 | 1 CX05_HUMAN | O75665 homo sapien |

ALIGNMENTS

| RESULT ID | MYHA_BOVIN | STANDARD | PRT | 1976 AA. |
|-----------|--|----------|-----|----------|
| AC | Q27991: | | | |
| DT | 16-OCT-2001 (Rel. 40, Created) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | |
| DE | Myosin heavy chain, nonmuscle type B (cellular myosin heavy chain, type B) (Nonmuscle myosin heavy chain-B) (NMHC-B). | | | |
| CN | MYH10. | | | |
| OS | Bos taurus (Bovine). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; | | | |
| OC | Bovidae; Bovinae; Bos. | | | |
| OX | NCBI_TaxID:9913; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Onara M., Ishiguro N., Shinagawa M.; | | | |
| RT | "Bos taurus nonmuscle myosin heavy chain B mRNA, complete cds."; | | | |
| RL | Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases. | | | |
| RN | [2] | | | |
| RP | SEQUENCE OF 204-302 FROM N.A. | | | |
| RC | TISSUE=Brain cortex; | | | |
| RX | MEDLINE=95301542; PubMed=7782316; | | | |
| RA | Itoh K., Adelstein R.S.; | | | |
| RT | "Neuronal cell expression of inserted isoforms of vertebrate nonmuscle myosin heavy chain II-B."; | | | |
| RL | J. Biol. Chem. 270:14533-14540(1995). | | | |
| CC | -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS, CAPING (BY SIMILARITY). | | | |
| CC | -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2). | | | |
| CC | -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS. | | | |
| CC | -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN. | | | |
| CC | -1- SIMILARITY: CONTAINS 1 IQ DOMAIN. | | | |
| CC | THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION - THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | |
| CC | EMBL: AB022023; BAA36494.1; - | | | |
| DR | EMBL: U15716; AAA87715.1; - | | | |
| DR | HSSP: P08799; IIVK | | | |
| DR | InterPro: IPR000048; IQ | | | |
| DR | InterPro: IPR004009; Myosin_N | | | |
| DR | InterPro: IPR002928; Myosin_Tail | | | |
| DR | InterPro: IPR002017; Spectrin | | | |
| DR | InterPro: IPR001609; myosin_head | | | |

| | | | | | |
|----|-------|-----|------|--------------|--------------------|
| 34 | 168.5 | 7.0 | 1941 | 1 MYH2_HUMAN | Q9UKX2 homo sapien |
| 35 | 168.5 | 7.0 | 2349 | 1 TPR_HUMAN | P12270 homo sapien |
| 36 | 168 | 7.0 | 697 | 1 MFPL_LYCES | P93203 lycopersico |
| 37 | 168 | 7.0 | 1972 | 1 MYHB_RABIT | P35748 oryctolagus |
| 38 | 167.5 | 7.0 | 727 | 1 MFPL_ARATH | Q91W85 arabidopsis |
| 39 | 167.5 | 7.0 | 1966 | 1 MYSB_CAEEL | P02566 caenorhabdi |
| 40 | 167.5 | 7.0 | 1978 | 1 MYHB_CHICK | P10587 gallus gall |
| 41 | 167.5 | 7.0 | 3911 | 1 AKA9_HUMAN | O99996 h-a-kinase |
| 42 | 166 | 6.9 | 1938 | 1 MYSD_CAEEL | P02567 caenorhabdi |
| 43 | 166 | 6.9 | 1972 | 1 MYHB_MOUSE | O08638 mus musculu |
| 44 | 165.5 | 6.9 | 579 | 1 G160_HUMAN | O08378 homo sapien |
| 45 | 165.5 | 6.9 | 848 | 1 MYSP_DIRIM | P13392 dirofilaria |

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AC Q99323;

DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, non-muscle (Zipper protein) (Myosin II).
 GN ZIP.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=90349606; PubMed=2117279;
 RA Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;
 RT "Complete sequence of the Drosophila nonmuscle myosin heavy-chain
 transcript: conserved sequences in the myosin tail and differential
 splicing in the 5' untranslated sequence."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).
 CC -1- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR
 CELLULARIZATION. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M35012; AAA28713.1; -.
 DR PIR: A36014; A36014.
 DR PIR: B36014; B36014.
 DR HSSP: P08799; IIMD.
 DR FLYBase: FBgn0005634; zip.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR004003; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 KW Myosin; Alternative splicing; Coiled coil; Actin-binding;
 KW ATP-binding; Calmodulin-binding.
 FT DOMAIN 1 829
 FT DOMAIN 830 859
 FT DOMAIN 886 2017
 FT NP_BIND 225 232
 FT DOMAIN 250 260
 FT DOMAIN 682 694
 FT DOMAIN 705 727
 FT DOMAIN 742 758
 FT DOMAIN 1303 1970
 FT DOMAIN 1971 2017
 FT VARSPIC 1 45
 FT SEQUENCE 2017 AA; 232016 MW; 73E3CB02BA8F2528 CRC64;

Query Match 7.9%; Score 190; DB 1; Length 2017;
 Best Local Similarity 23.8%; Pred. No. 0.025;
 Matches 88; Conservative 61; Mismatches 145; Indels 76; Gaps 14;

QY 56 KTIINKLFFDLAEEENVLDAE--FLKNELDSVKAQLSQKDREKRSQ-----AI 103
 Db 1263 KTVLEK-----AKGTLEENADLATELRSVSSNQENDRRKQAESQIAELQVLAE 1314
 QY 104 IDTLRDTLEER---NATVESLONALKAEKLCSTLKKQKFEQRQDEFKQAREEAAHR 158
 Db 1315 IERARSELQEKCTKLQOEAENITQMLEAEVLKSAVKSASNMESQLEAOQLLEEBTRQ 1374
 QY 159 ---LKCKMKTMQEQLELLQSORSEVEEMIRDMVGQSAVEDLAVYCSLKREYNLKEAR 215
 Db 1375 KLGSSKRLQTESKEALQEOLEDEEDAKRNY---EKRLAVTTQMOEIKKRAEDADLA 1431
 QY 216 KATGELADRLKDLVSSRSKLTL--NTELDQAKLELSAQKD---LQADQETSLR 268
 Db 1432 KELEEGKRLNKDLEALERQVKELAQNDRLDKRKKIQSLSEATTELEQRKVLLE 1491
 QY 269 KKSDDPGNLEPASATNETVSRLVFESSPAPYEMNPRLOHPFGEIDLTFTDVNTPPT 328
 Db 1492 KKQK---NFDKILAEKALISEQIAQERDPTREAREKETVLVSRELDEAFD----- 1541
 QY 329 QTSQSQCPLPKLCLERARSPQNVL-----KVKVAKP-----ESQIS-LGQ 372
 Db 1542 -----KIEDLENKRTLQNELDLANTQGTADKNVHELEKARALESLAEIKAQ 1591
 QY 373 RCVGELDEEL 382
 Db 1592 N-ELEDDL 1599
 RESULT 4
 MYSC.CHICK STANDARD; PRT; 1102 AA.
 AC P29616;
 ID MYSC.CHICK
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, cardiac muscle isoform (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE OF 65-1102 FROM N.A.
 RC STRAIN=BROILER BREEDERS; TISSUE=Heart;
 RX MEDLINE=92130260; PubMed=1774788;
 RA Stewart A.F.R., Camoretti-Mercado B., Perlman D., Gupta M.,
 RA Jakovic S., Zak R.;
 RT "Structural and phylogenetic analysis of the chicken ventricular
 RT myosin heavy chain rod."
 RL J Mol. Evol. 33:357-366(1991).
 RN [2]
 RP SEQUENCE OF 1-259.
 RC TISSUE=Heart;
 RX MEDLINE=93039740; PubMed=1418675;
 RA Metanabe B.;
 RT "Amino-acid sequence of the short subfragment-2 in adult chicken
 RT cardiac muscle myosin."
 RL Biol. Chem. Hoppe-Seyler 373:1045-1054(1992).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- TISSUE SPECIFICITY: VENTRICULAR MUSCLE AND TRACES LEVELS IN THE
 CC DEVELOPMENT AND REGENERATION FOLLOWING FREEZE INJURY.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES.
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED

CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,
CC WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE
CC CONSERVED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59552; CAA42130.1; -
DR HSSP: P03069; 22TA.
DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PF01576; Myosin_tail; 1.
KW Myosin; Muscle protein; coiled coil; thick filament; Actin-binding;
KW ATP-binding; Multigene family.
KW NON_TER 1
FT DOMAIN <1 1102 COILED COIL (POTENTIAL).
FT CONFLICT 65 65 A -> T (IN REF. 1).
FT CONFLICT 92 92 V -> A (IN REF. 1).
FT CONFLICT 113 114 CS -> PAL (IN REF. 1).
FT CONFLICT 135 135 H -> Q (IN REF. 1).
FT CONFLICT 185 185 T -> R (IN REF. 1).
SQ SEQUENCE 1102 AA; 128008 MW; 2293668D049825DC CRC64;

Query Match 7.8%; Score 185.5; DB 1; Length 1102;
Best Local Similarity 20.2%; Pred. No. 0.022;
Matches 77; Conservative 82; Mismatches 167; Indels 55; Gaps 10;

QY 59 INKLEFDLQAEENVL--DAEFLKELNLDVSKAQLSOKREKRDQAIDTLRDLLEENA 116
DB 432 MTRLMNDLTQTKYLOSENGEVR-QLEEKESLISQSRKSTFTQQLERQLEERTK 490
QY 117 TVESIQNALAKEMICSTLKKRMKFLRQDETRQAREARLRCKMKMTQEIQLLSQ 176
DB 491 SKMLAHALQARHDCDLRQYEEEDQAKALDRAISKGAFAVQMTKTKETDAIQTE 550
QY 177 RSE--VEEMITMGVGSVAEQLAVYCVSLKEKENK-----EARKATGELADRLK 226
DB 551 ELDEAKKKRLRLQEAELAEANAKCSLEKAKHRLQNEQEDMDIMLEKANSAAISDK 610
QY 227 KDIVSSRSKLTN--TELQAKLELRSAQKDLQSAQDELTSLRKSDDPGULERASA 283
DB 611 K---QKGFDTIINDMKOKTEESQAELEASQKARSUSTELFKLMAVEETLDHLETLKR 666
QY 284 TNETVSRVLESPARVEKMNPRLLHQPFGEIDINTTFDVNTPTQTSGSQHCLPKKLT 343
DB 667 ENKMLQELISDLTNOISEGNKHL-----EIE-----KVKQVDEKSEVQDAL 710
QY 344 ERANSPQONVLLKVKHAKSKPESQSLSGQRCVGLDEDELAGAPFLFRNAVLAGKQPNRT 403
DB 711 EEAEGALEHEESKTLRQLSELQKADFERKLAEDDEMQNI-----RRNOQRT 759
QY 404 -----TAESRSSTDVRI 416
DB 760 IDSLQSTLDSEARSRNEAIRL 780

RESULT 5
MYH3_RAT STANDARD: PRT; 1940 AA.
ID MYH3_RAT P12847;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87060988; PubMed=3783701;
RA Strehler E.E., Strehler-Page M.-A., Perriard J.C., Perlasamy M.,
RA Nadal-Ginard B.;
RT "Complete nucleotide and encoded amino acid sequence of a mammalian
RT myosin heavy chain gene. Evidence against intron-dependent evolution
RT of the rod.";
RL J. Mol. Biol. 190:291-317(1986).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES.
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04267; CAA27817.1; -
DR PIR: A24922; A24922.
DR HSSP: P08799; LMND.
DR InterPro: IPR000048; IO.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IO; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IO; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IO; 1.
KW Myosin; Muscle protein; Coiled coil; thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IO.
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 656 678 ACTIN-BINDING.
FT DOMAIN 758 772 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 696 696 ALKYLATION (SH-1).
FT MOD_RES 706 706 ALKYLATION (SH-2).
SQ SEQUENCE 1940 AA; 223857 MW; B5D546A596E5A696 CRC64;

Query Match 7.8%; Score 185.5; DB 1; Length 1940;
Best Local Similarity 19.5%; Pred. No. 0.041;
Matches 74; Conservative 84; Mismatches 180; Indels 41; Gaps 7;

QY 54 GKTIINKLEFDLQAEENVL--DAEFLKELNLDVSKAQLSOKREKRDQAIDTLRDLLE 112

```

Db 1262 GKNEETORSISELTQKSRLOTEAGELSRQLEEKESIVSOLSRKQAFQTOIEELKROLE 1321
QY 113 ERNATVESLONALKAMLCSTLKKOMKFLQRODETKQAREFHRLKCKMKTMEQIELL 172
Db 1322 EENKAKNALHAAQSSHHDDLLREQYEEDEQGAELQRLAKLSKNSVAAWRTKYETDAI 1381
QY 173 -----LOSORESEVEMIRDMGVGOSAVEOLAVYCVSLK-----EYENIKEARKATGE 220
Db 1382 QRTEELEBAKKLAQRQD---SEQYEVAVNACASLEKTKQROLOGEVEDLMDVVEANS 1438
QY 221 LADRLKDLVSSRSKLTTELTDQAKLELSAQKDLQASQDOETTSLRKSDPPGNLEP 280
Db 1439 LAALDKQKNPFKVLAEWKTKCEEQAELEALKESSRLSTELFLKLNAYEALDQLET 1498
QY 281 ASATNEVSVLVEESPAPVEMNPRHLQPPFGDEIDLTNTFDVNTPTPTQSSGHCLPK- 339
Db 1499 VKRENNKLEQELDLTEQIAENKSKSHE-----LEKSRKQMELEKA 1539
QY 340 --KLCLEKARSPMONVLYKRVKHKVSKPESQSLSGQRCVGELEDELAGAFLEIRNAVIGQ 397
Db 1540 DIGMALEEAALAEHEHAKLRLQLELTQYKSEIDRKIAKDEIDELQKNYQRTV---E 1596
QY 398 KQPNRTTAESRSSTDVRI 416
Db 1597 TMOGALDAEVRSRNEAIRL 1615

```

RESULT 6

MYHA_RAT STANDARD; PRT; 1976 AA.

AC 09JLTO;

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Myosin heavy chain, nonmuscle type B (Cellular myosin heavy chain,

type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).

GN MYH10.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;

RA Yam J.W.P., Chan K.W., Hsiao W.L.W.;

RT Suppression of transverse phenotype by overexpression of nonmuscle

myosin heavy chain-B in p53val135-transformed rat fibroblast.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,

CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND

CAPPING.

-!- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY

CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC-2).

-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING

CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,

CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

-!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

-!- SIMILARITY: CONTAINS 1 IQ DOMAIN.

OS This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

CC EMBL: AF139055; AF61445.1; -

DR InterPro: IPR000048; IQ.

DR InterPro: IPR004009; Myosin_N.

DR InterPro: IPR002928; Myosin_tail.

```

DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00063; myosin_head.1.
DR Pfam: PF02736; Myosin_N.1.
DR Pfam: PF01576; Myosin_tail.1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head.1.
DR SMART: SM00015; IQ.1.
DR SMART: SM00242; MYSC.1.
DR PROSITE: PS50096; IQ.1.
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Alkylation; Multigene family.
FT DOMAIN 1 785
FT MYOSIN HEAD-LIKE.
FT DOMAIN 786 815
FT COILED COIL (POTENTIAL).
FT NP_BIND 178 185
FT ATP (POTENTIAL).
FT MOD_RES 701 701
FT ALKYLATION (SH-2) (POTENTIAL).
FT MOD_RES 711 711
FT ALKYLATION (SH-2) (POTENTIAL).
SQ SEQUENCE 1976 AA; 228963 MW; E32708BF9BF2B470 CRC64;

```

Query Match 7.88; Score 185.5; DB 1; Length 1976;
Best Local Similarity 22.18; Pred. No. 0.042;
Matches 77; Conservative 77; Mismatches 139; Indels 55; Gaps 10;

```

QY 68 QEEENVLDAEFLKNEIDSVKQAQLSQKDRKDSQAIIIDTLRDLTEERNATVESLONLTK 127
Db 1505 QNKOLRADMEDLMSSKSDVGNVHELEKSKRALEQYVEEMTQLEDELEQATEDAKLR 1564
QY 128 AEMLCSTLKKOMKFLQRODETKQAREHRLKCKMKTMEQIELLOSORSE----- 179
Db 1565 LEVWQMKAQOFERDQTRDE---QNEKKRL--LKQVHELAELEDERKQRLAVASK 1619
QY 180 --VEEMTRDMGVGOSA-----VEQLAVYCVSLKREYENLKKARKATGLADRLKKDL 229
Db 1620 KKMEDIKDLDEAQLQAEANKARDEVIAKQRLQAKMDYORELEBARASRDEIFAQSK-- 1677
QY 230 VSSRSKLTTELTDQAKLELSAQKDLQASQDOETTSLRKSDPPGNLEPASTNTGVS 289
Db 1678 --SEKRLKLSLEAELTQLOEELASSERARRHAEQDRDEL--ADEIANSASGKSLADEKR 1732
QY 290 RL-----VESAPVEMNPRHLQPPFGDEIDLTNTFDVNTPTPTQSSGHCLPK 339
Db 1733 RLEARIQAEELEEBESQNNMELNDRPK-----ITLQVDLTLELAARSAQAQ 1782
QY 340 ----KLCLEKARSPMONVLYKRVKHKVSKPESQSLSG--QRCVGELEDEEL 382
Db 1783 SDNARQOLEQRQNKELKAKLOLEEGAVKSKFATSALEAKIGOLEEQL 1830

```

RESULT 7

MYSL_DICDI STANDARD; PRT; 2116 AA.

AC P08799;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin II heavy chain, non muscle.

GN MHCA.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

OX NCBI_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=87092266; PubMed=3540939;

RT "Conserved protein domains in a myosin heavy chain gene from

Dictyostelium discoideum";

RU Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).

RN [2]

RP PHOSPHORYLATION SITES, AND MUTAGENESIS.

RC STRAIN=AX2;

MDL=90353583; PubMed=2387408;

RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,
 RA Gerisch G.;
 RT "Replacement of threonine residues by serine and alanine in a
 RT phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
 RL FEBS Lett. 269:239-243(1990).
 RN [3].
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=88112226; PubMed=2828113;
 RA Magle G., Noegel A., Scheel J., Gerisch G.;
 RT "Phosphorylation of threonine residues on cloned fragments of the
 RT Dictyostelium myosin heavy chain.";
 RL FEBS Lett. 227:71-75(1988).
 RN [4].
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345066; PubMed=7619795;
 RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
 RA Rayment I.;
 RT "X-ray structures of the myosin motor domain of Dictyostelium
 RT discoidium complexed with MgADP. Bepx and MgADP.ALPA4-";
 RL Biochemistry 34:8960-8972(1995).
 RN [5].
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345067; PubMed=7619796;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
 RT truncated head of Dictyostelium discoidium myosin to 2.7-A
 RT resolution";
 RL Biochemistry 34:8973-8981(1995).
 RN [6].
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=96206189; PubMed=8611530;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the
 RT Dictyostelium discoidium myosin motor domain to 1.9-A resolution.";
 RL Biochemistry 35:5404-5417(1996).
 RN [7].
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
 RX MEDLINE=97452580; PubMed=9305951;
 RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
 RT "X-ray structures of the MgADP, MgATPgammass, and MgAMPNP complexes
 RT of the Dictyostelium discoidium myosin motor domain.";
 RL Biochemistry 36:11619-11628(1997).
 RN [8].
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=98070605; PubMed=9405148;
 RA Bauer C.B., Kuhlman P.A., Bagnath C.R., Rayment I.;
 RT "X-ray crystal structure and solution fluorescence characterization
 RT of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
 RT Dictyostelium discoidium myosin motor domain.";
 RL J. Mol. Biol. 274:394-407(1997).
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
 CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
 CC SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
 CC INTO FILAMENTS. HEAVYMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
 CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 CC (MLC-2).
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
 CC CORTEX.
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER
 CC SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
 CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
 CC ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
 CC POSITION (688).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14628; AAA33227.1; -
 DR PIR: A26655; A26655.
 DR PIR: S00250; S00250.
 DR PDB: 1MMA; 03-DEC-97.
 DR PDB: 1MMD; 17-AUG-96.
 DR PDB: 1MMG; 03-DEC-97.
 DR PDB: 1MMN; 03-DEC-97.
 DR PDB: 1MND; 17-AUG-96.
 DR PDB: 1MNE; 17-AUG-96.
 DR PDB: 1VOM; 23-DEC-96.
 DR PDB: 1LVR; 28-JAN-98.
 DR DictyDb; DD01008; mhca.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam; PF00612; IQ; 2.
 DR Pfam; PF00063; myosin_head; 1.
 DR Pfam; PF02736; Myosin_N; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IO; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS50096; IO; 1.
 KW Myosin: Coiled coil; Actin-binding; ATP-binding; 3D-structure;
 KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
 FT DOMAIN 1 761 MYOSIN HEAD-LIKE.
 FT DOMAIN 762 791 IO.
 FT DOMAIN 817 2116 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP.
 FT DOMAIN 638 660 ACTIN-BINDING.
 FT DOMAIN 738 752 ACTIN-BINDING.
 FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).
 FT MOD_RES 678 678 ALKYLATION (SH-1).
 FT MOD_RES 1823 1823 PHOSPHORYLATION (BY MHCK).
 FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MHCK).
 FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
 FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
 SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770B1EE56A1 CRC64;
 Query Match 7.88; Score 185.5; DB 1; Length 2116;
 Best Local Similarity 25.14; Pred. No. 0.045;
 Matches 82; Conservative 60; Mismatches 108; Indels 77; Gaps 15;
 QY 60 NKLFDLAQEEENVLDLAE-FLKKNLDSVKAQLQSKD--REKRPDQAIIDRLRPTLEERN 115
 DB 852 DKLEKSLKTESNVLDLQROKAKETUKAYDSKDLAKQKLELRVEDMESELDKK 911
 QY 116 ATVESLONALNKAKMLCTLKQKMFLEORDETQKQAREBAHRLECKMKTEQIELLLOS 175
 DB 912 LALENLQNKRSVE-----EKVRDLKEELQEEQKLRNLETKAKKYVE----- 953
 QY 176 QRSVEEKIRDMGVGS-ANEQLAVYCVSLKKEYNLEKARKATGELADRLKLDVSSRS 234
 DB 954 --EELEEKRR-VNDGOSIDTISRLKIRDELQKEVEELTES-----PSESKDKGVLEKT 1004
 QY 235 KIKLTNLELQAKLELSAQKDLQSAQOEITSLKKSDDPPGNLEPASTNTEYSLVFE 294
 DB 1005 RVR-LQSELDLVLRDSETD---KSELRLQKKLELEIKQVEALAA-ETAKLAQE 1058
 QY 295 SPAPVEKMNRLHPQPGDEIDLNTTFPDVNTPPQTGSGHCLPKKLCLEBARSPMONVL 354
 DB 1059 A-----ANKKLG-----GETTELNEKF--NSEVTARSN-----VEKSKTLESOL 1096
 QY 355 KYHKVSKPESQSLSGGRCVGLDEE 381
 I: |||||

Db 1097 VAWN-----NEEDEE 1106

RESULT 8

MYH3_HUMAN STANDARD; PRT; 1940 AA.

AC P11055; Q15492; (Rel. 11, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, fast skeletal muscle, embryonic (muscle embryonic myosin heavy chain) (SMHCE).

GN MYH3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP MEDLINE=99263803; PubMed=2726495;

RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H.,

RT Rubinstein N.A., Kelly A.M., Sarkar S.,

RT "Nucleotide sequence of full length human embryonic myosin heavy chain cDNA."

RL Nucleic Acids Res. 17:3591-3592(1989).

RN [2]

RP MEDLINE=90033296; PubMed=2805546;

RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L.,

RT Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.,

RT "Human embryonic myosin heavy chain cDNA. Interspecies sequence conservation of the myosin rod, chromosomal locus and isoform specific transcription of the gene."

RL FEBS Lett. 256:21-28(1989).

RN [3]

RP MEDLINE=856-1940 FROM N.A.

RC TISSUE=Skeletal muscle;

RX MEDLINE=90235862; PubMed=1691980;

RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goede H.W.,

RT Arnold H.H.,

RT "Identification of three developmentally controlled isoforms of human myosin heavy chains."

RL Eur. J. Biochem. 189:55-65(1990).

RN [4]

RP MEDLINE=856-1940 FROM N.A.

RA Karsch-Mizrachi I., Travis M., Blau H., Levinand L.A.;

RT "Expression and DNA sequence analysis of a human embryonic skeletal muscle myosin heavy chain gene."

RL Nucleic Acids Res. 17:6167-6179(1989).

CC -1- FUNCTION: MUSCLE CONTRACTION.

CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

CC -1- DEVELOPMENTAL STAGE: ABUNDANTLY PRESENT IN FETAL SKELETAL MUSCLE AND NOT PRESENT OR BARELY DETECTABLE IN HEART AND ADULT SKELETAL MUSCLE.

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PFM: TWO CYSTEINE RESIDUES IN THE SI DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X13988; CAA32167.1; -

DR EMBL: X13100; CAA31492.1; -

DR EMBL: X51593; CAA35942.1; -

DR EMBL: X15696; CAA33731.1; -

DR PIR: S04090; S04090.

DR HSSP: P08799; 1MD.

DR MIM: 160720; -

DR InterPro: IPR000048; IQ.

DR InterPro: IPR004009; Myosin_N.

DR InterPro: IPR002928; Myosin_tail.

DR InterPro: IPR001609; myosin_head.

DR Pfam: PF00612; IQ; 2.

DR Pfam: PF00063; myosin_head; 1.

DR Pfam: PF02736; Myosin_N; 1.

DR Pfam: PF01576; Myosin_tail; 1.

DR PRINTS: PR00193; MYOSINHEAVY.

DR ProDom: PD000355; myosin_head; 1.

DR SMART: SM00015; IQ; 1.

DR SMART: SM00242; MYSC; 1.

DR PROSITE: PS50096; IQ; 1.

DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation; Multigene family.

KW MYOSIN HEAD-LIKE.

FT DOMAIN 1 781

FT MOD_RES 782 811

FT NP_BIND 840 1933

FT NP_BIND 179 186

FT DOMAIN 656 678

FT MOD_RES 758 772

FT MOD_RES 130 130

FT MOD_RES 696 696

FT MOD_RES 706 706

FT CONFLICT 1331 1331

FT CONFLICT 1391 1392

FT CONFLICT 1608 1609

FT CONFLICT 1663 1664

SO SEQUENCE 1940 AA; 224035 MW; 43CA58C6A4BA1253 CRC64;

Query Match 7.7%; Score 184.5; DB 1; Length 1940;

Best Local Similarity 19.5%; Pred. No. 0.046;

Matches 74; Conservative 85; Mismatches 179; Indels 41; Gaps 7;

QY 54 GKKTINKLFDLAQEEENV-LDEFLKNELDVSKAQLSOKDRKRSQATIDRLDRL 112

DB 1262 GKNEIORSLELTQKSRLOTGEGLSRLEKEESIYSQLSRKSQAFQTQTEELROLE 1321

QY 113 ERNATVESIQNALNKAEMLCSTLKQMKFLERODETQKAREEHRILCKKKTMEQIELL 172

DB 1322 EENKAKNALAHALQSSRRDCDLRQVEEEOGAELQALSKANSEVAQRTYVEDAI 1381

QY 173 -----LOSQSEVEEMIRDMGVGQSAVQQLAVYCVSLKK-----EYENIKKARKATGE 220

DB 1382 QRTLELEAKKKLQRLQD---SEQVAVNAKCASEKTKQRLQGEVEDLMDIVERANS 1438

QY 221 LADRLKQVLVSRSKGLKTLNLEQAKLELSAQKDLQSAQDETSLSKKSDDPPGNLEP 280

DB 1439 LAAALDKQKRFNDKYLAEWKTRCEQSLKESKSTLEFLKLNAYEALDQLET 1498

QY 281 ASATNETVSRVLFVSPAPVEMNNPRLQPPGDEIDLITPDVMTPPQTSQSQCIPK- 339

DB 1499 VKREKNLEQEIADLTEDIAENGKTIHE-----LEKSRKQIELEKA 1539

QY 340 --KICLERARSPMONVLKVKHVKSPESQSLSGQRCVGLDEELAGAFPLIRNAVLCQ 397

DB 1540 DIQALAEAEALAEHBEAKIIRIOLLETQVKESEIDRKIAERDEDEQLKRNVOFTV--E 1596

QY 398 KQPNNTAESRSSRDVARI 416


```

Db 1597 TMOSALDAEVRSRNEAIRL.1615

RESULT 9
MYS_AEOIR STANDARD; PRT; 1938 AA.
ID_MYS_AEOIR
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, striated muscle.
OS Aequipecten irradians (Bay scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
OC Pectinoidea; Pectinidae; Argopecten.
OX NCBI_TaxID=31199;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=92011595; PubMed=1917970;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Nucleotide sequence of full length cDNA for a scallop striated
muscle myosin heavy chain."
RL Nucleic Acids Res. 18:7158-7158(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=91088319; PubMed=2263488;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Nucleotide sequence of full length cDNA for a scallop striated
muscle myosin heavy chain."
RL Nucleic Acids Res. 18:7158-7158(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX MEDLINE=94173332; PubMed=8127365;
RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,
RA Szent-Gyorgyi A.G., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2.8-A
resolution."
RL Nature 368:306-312(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
RX MEDLINE=96419133; PubMed=8805510;
RA Houdusse A., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2-A
resolution: implications for regulation."
RL Structure 4:21-32(1996).
CC - FUNCTION: MUSCLE CONTRACTION.
CC - ACTIVITY: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC - SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC - SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC - SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC - SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X55714; CA39247.1; -.
DR PIR: S13557; S13557.
DR PIR: A40997; A40997.
DR PDB: 1SCM; 30-APR-94.
DR PDB: 1MDC; 11-JUL-96.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.

```

```

DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW MYOSIN: Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
FT DOMAIN 1 777
FT DOMAIN 778 805
FT DOMAIN 836 1938
FT DOMAIN 836 1938
FT NP_BIND 176 183
FT MOD_RES 693 693
FT MOD_RES 703 703
FT HELIX 778 821
FT TURN 822 823
FT TURN 825 833
FT TURN 834 835
SQ SEQUENCE 1938 AA; 222821 MW; A5CCEA127D1A4896 CRC64;

```

```

Query Match 7.68; Score 183; DB 1; Length 1938;
Best Local Similarity 23.0%; Pred. No. 0.054;
Matches 102; Conservative 67; Mismatches 128; Indels 146; Gaps 20;

```

```

QY 55 KKTINKLFDP-----LAQEEENVDAEFLKNEIDSVKADL-----SQ 92
DB 903 EKLMOQADFEFSQIKLEERLDEEDAAADLEGTKKMEADNANLKKDIDGLENTLOKAE 962
QY 93 KDKERKDSQALITDYL-----RDTLEERN-ATVESIQ-----NALNKA 128
DB 963 QDKAHKDNQ--ISTLGEISQDDEHICKLNKKEKALKEANKKTSDSLQAEEDKCNHLNK 1019
QY 129 EMLCSTLKKQKFLFQRODETKQAREENHRLKCKMK-----TMCOTELLLOSREVEEM 183
DB 1020 --LKAQLEQALDELEDNLEREKKVRGVEKAKRVEDDLKSTQENVEDLERVKR-ELEEN 1076
QY 184 IR-----DMGQSAVEQLAVCYSLKKEYENLKARATGELMDRLKKDLVSS 232
DB 1077 VRKKEAEISSLNSKLEBDQNLVSOLOKRIKELQRIEELERLEAERNARAKVKQRAEL 1136
QY 233 RSKLKTINTLELDOA-----KLELRSAQKDLQSA-----DOETSLRKSDDP 274
DB 1137 NRELEIEGERLDEAGATSAQIELNKKREAEELKIRRDLEEASQHQHQAQISALKKKHQD 1195
QY 275 PGNLEPASATREYVSRVFESPAVEEMNRLHOPPGDEIDLTNTTVDVNTPPPTGSGQ 334
DB 1196 -----AANEMADQV-----DQLQKVSKLEK-----DKKDLKREMD-----DLESQMT 1233
QY 335 HCLPKKLCLEPARSPQNVNLKKVKKVSKPSQSLS-----LGGORCGELEDLQAGFPL 388
DB 1234 HNMKNKCCSEK-----VKQF-----ESQMSDLNARLEDSQRSINEL----- 1270
QY 389 FIRNAVIGOKOPNRTTAESRSST 411
DB 1271 -----QSQKSRLOAENSDLT 1285

RESULT 10
MYSU_RABIT STANDARD; PRT; 501 AA.
ID_MYSU_RABIT
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Myosin heavy chain, embryonic smooth muscle isoform (Fragment).

```



```

DR Prodom: PD000355; myosin_head: 1.
DR SMART: SM00015; IQ: 1.
DR SMART: SM00242; MYSC: 1.
DR PROSITE: PS50096; IQ: 1.
KW Myosin: Muscle protein; Coiled coil: Thick filament; Actin-binding:
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation:
KW Multigene family.
FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
FT DOMAIN 785 814 COILED COIL (POTENTIAL).
FT NP_BIND 843 1938 ATP (POTENTIAL).
FT DOMAIN 179 186 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 761 775 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 130 130 METHYLATION (SH-1) (POTENTIAL).
FT MOD_RES 699 699 ALKYLATION (SH-2) (POTENTIAL).
FT MOD_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).
SO SEQUENCE 1938 AA: 223678 MW: 1160006416381CDS CRC64:

Query Match 7.5% Score 180; DB 1: Length 1938;
Best Local Similarity 21.1% Pred. No. 0.076;
Matches 84; Conservative 67; Mismatches 122; Indels 126; Gaps 13;

QY 68 QEEENVL-----DAEFLKNEIDSKVAQLSQDKREK-----RDSQALITDRLRTLEERN 115
DQ 1296 EERKESLSIQLTRKQALTOOLEELKROMEEETKAKNAMAHALOSSRHDCDLREQYEEQ 1355
QY 116 ATVESIQNALNKAE-----MLCSTLKQMKFLEORODETKQAREEAHRLK 160
DQ 1356 EAKAELOALSKANSEVAMQMTKYETDAIQRTEELEAKKLAQLOAEKETEYANS-K 1414
QY 161 CKMKTMEOIELLOSQSEVEEMIRDMGVGSAVEQOLAVCVSLKKEYENLKEARKATGE 220
DQ 1415 C-----ASLETKQTLQEEVEDLMDL-----ERSHTACATLIDKKORNDKYLAEWKQ 1462
QY 221 LADRLKQVSSRSKLTQNTL-----DQAKLELRSAQKQDQSDQETSRLK 269
DQ 1463 KLDSEQAELKAEKESKSLSTELFKRNAYEEVDQLE-TLRREKKNLQ-----EETSDLTE 1518
QY 270 KSDPPGMLPEPASATNETVSRLVESPAFVEMMNPRLQHPFGDEIDLTFTVDVTPPTQ 329
DQ 1519 QIAETGKNLQAEKTKKVEQ-----EKSDL----- 1544
QY 330 TSSGSHCLPKKLCLEPARSPQONVKKYKVKSPBSQLSLGQRCVGLDELAGAPFLF 389
DQ 1545 -----QVLEEEVGSLEHESKSLRVQLSEGVKSELDRKVIKDEBI----- 1587
QY 390 IRNAVLOKQKOPNRT-----AESRSSDVVRI 416
DQ 1588 -----EOLKRSQRAALQSVLDAELIRSRDMLRL 1618

RESULT 12
MYSA_CAEEL STANDARD; PRT: 1969 AA.
AC P1284;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain A (MHC A).
GN MYO-3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=89178677; PubMed=2926820;
RA Ddbb N.J., Maruyama I.N., Krause M., Karn J.;
RT "Sequence analysis of the complete Caenorhabditis elegans myosin
heavy chain gene family.";
RL J. Mol. Biol. 205:603-613(1989).

```

```

CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
CC C. ELEGANS.
CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
DR EMBL: X08067; CA30856.1; -.
DR PIR: S02771; S02771.
DR HSP: P08799; ILVK.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ: 1.
DR Pfam: PF00063; myosin_head: 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head: 1.
DR SMART: SM00015; IQ: 1.
DR SMART: SM00242; MYSC: 1.
DR PROSITE: PS50096; IQ: 1.
KW Myosin: Muscle protein; Coiled coil: Thick filament; Actin-binding:
KW ATP-binding; Methylation; Alkylation; Multigene family.
KW DOMAIN 1 793 MYOSIN HEAD-LIKE.
FT DOMAIN 794 823 IQ.
FT DOMAIN 857 1969 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP.
FT DOMAIN 667 689 ACTIN-BINDING.
FT DOMAIN 770 784 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 707 707 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 717 717 ALKYLATION (SH-2) (POTENTIAL).
SO SEQUENCE 1969 AA: 225509 MW: 64577BBAE7EAD80A CRC64:

Query Match 7.5% Score 180; DB 1: Length 1969;
Best Local Similarity 21.2% Pred. No. 0.077;
Matches 126; Conservative 85; Mismatches 203; Indels 180; Gaps 24;

QY 31 LQCLIQFETAPSRPCQCRIOVGKTTINKLPFDLQEEENVLDAEFLKNEIDSKVAQ 90
DQ 872 LEEAVQGETARSQLSEQVADLVEEK---NALPISLETETKANLADAERNEKLNQATL 928
QY 91 SQK-----DREKDSQAIIPT-----LRDTLEE---RNATV 118
DQ 929 ESKLSIDITGOLEDQERNEDLARQKKTKDQSLTKKHVDQLSLKAKQDEKSRDHN 988
QY 119 ESLQNALNKAEMLCSTLKQMKFLEORODETKQAREEAHRLKCKMKTMEOIELLOSQ 178

```


GN US01 OR INT1 OR YDL058W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RX MEDLINE=91185402; PubMed=2010462;
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
 RA Yamasaki M.,
 RT "A cytoskeleton-related gene, us01, is required for intracellular
 RT protein transport in Saccharomyces cerevisiae."
 RL J. Cell Biol. 113:245-260(1991).
 RN [2]
 RP SEQUENCE OF 782-1790 FROM N.A.
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
 RA Kendrick K.E.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-8 FROM N.A.
 RA Bai Y., Symington L.S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
 CC ER AND THE GOLGI COMPLEX.
 CC -1- DOMAIN: THE RODLIKE TAIL. SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
 CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X54378; CA38253.1; -;
 DR EMBL: L03188; AAB00143.1; -;
 DR EMBL: U53668; AAB66559.1; -;
 DR PIR: A38455; A38455.
 DR HSSP: P80220; IDIP.
 DR SGD: S0002216; USO1.
 DR InterPro: IPR002017; Spectrin.
 KW Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
 FT DOMAIN 1 724 GLOBULAR HEAD.
 FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
 FT DOMAIN 991 1790 CHARGED (HYPER-HYDROPHILIC).
 FT DOMAIN 1172 1786 DISPENSABLE FOR THE PROTEIN FUNCTION.
 FT CONFLICT 847 847 ASP/GLU-RICH (ACIDIC).
 FT CONFLICT 924 924 G -> E (IN REF. 2).
 FT CONFLICT 1253 1253 V -> K (IN REF. 2).
 FT CONFLICT 1319 1319 V -> I (IN REF. 2).
 FT CONFLICT 1461 1461 I -> V (IN REF. 2).
 FT CONFLICT 1581 1581 N -> S (IN REF. 2).
 FT CONFLICT 1600 1600 G -> S (IN REF. 2).
 FT CONFLICT 1661 1661 I -> V (IN REF. 2).
 FT CONFLICT 1772 1772 R -> S (IN REF. 2).
 FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).
 SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9F04818 CRC64;

Query Match 7.4%; Score 178; DB 1; Length 1790;
 Best Local Similarity 18.1%; Pred. No. 0.087;
 Matches 81; Conservative 110; Mismatches 177; Indels 80; Gaps 13;

QY 56 KTIINKLFEDLAEEVNLDA-EPLKNE---LDSVKAQLSQKDRKRDSDQAIITDLRDT 110
 DB 965 KSLANN-YKDMQANESLIRAVESKNSSIOISLNKIDMSQSEKENFOIERGSTEKN 1023

QY 111 LEERNATVESLQNA-----LNKAEMLCSTLKKOMKFLERODETKOAREE-AHRLCKMKYT 165
 DB 1024 IEOLKRTISDLEQTKREILISDSKDEYESQISLLEKLETATTAENDENKKSILTKYT 1093
 QY 166 MEQIELLLQSQRSEVEEMIRDMGVQSAVEQLAVYCSLKEYNELKARKATGELADRL 225
 DB 1084 REELEAEILAAVKKNLNELETLETSEKALKE-----VKNEBHLKEEKIQLEKATYT 1136
 QY 226 KKDLYSSRSKKTINTLEIDAKKLELRSNQK-----DLQSAQOETISL 267
 DB 1137 KOQINSLRANLESLEKHEHDLAQLKRYEYIANKEROYNEISQLNDEITSTQOENESI 1196
 QY 268 RKSDDPPGNLEPASATNETFVSRLVFPESPAPVEMNRLHOPPGDEIDLTTTVDVNTMP 327
 DB 1197 KKKHDELEGEYKAMKSTSEBSNL---KKSIDALNIQIK- 1234
 QY 328 TQTSQSQHCLPKLCLERAR--SPQNVLYKKVHVSRPESQSLSGQRCVGEIDELAG 385
 DB 1235 -----LKKNETNEASLLLESIKSVESETVYKIKELQDECNF-KEKEVSELEDKLKAS 1284
 QY 386 PPLIRNAVLGQKQPNNTTASBSTDVVRIGFGGLGRKTFIOPRDTTITRPVYKSK- 444
 DB 1285 EDKNSKYLEL-QKSEKIKETLDAKTTLEKIQLEKITNLNKAKESESELRLKRTSSEE 1343
 QY 445 -----AKSKORIKRTVSSASQPKL 464
 DB 1344 RKNMEQLEKLNKNIQIKNAFKEKRL 1371
 RESULT 15
 ID YD86.SCHPO STANDARD; PRT; 1957 AA.
 AC 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 227.8 kDa protein ClF3.06C in chromosome I.
 GN SPAC1P3.06C.
 OS Schizosaccharomyces pombe (fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomyces.
 CX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z70690; CAA94624.1; -;
 DR Hypothetical protein.
 KW Hypothetical protein.
 SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

Query Match 7.4%; Score 177.5; DB 1; Length 1957;
 Best Local Similarity 20.6%; Pred. No. 0.1;
 Matches 74; Conservative 70; Mismatches 135; Indels 81; Gaps 9;

QY 70 EENVYDAEFLKNELDLSKQALDSQKDRKRDSDQAIOTLPTLEERNATVESLQNALKAE 129
 DB 1386 EDNOLATNKLKNDQDHLNDQIRUKEDVLYKRESLITLSLESLSNQROKESLIDAKNELE 1445
 QY 130 -MLCSTLKKOMKFLERQ-----DETQAREBAHRLCKMKMTMEQIELLLQSQRSEVE 181

Db 1446 HMLDTSRKNSSLMKIESINSSLDKSFELASAVEKLGALQKLHSELSIMENIKSOLQ 1505
 QY 182 EMIRDMGVGASAVEQLAVVCYLKREXEN-----LKEARKATG 219
 Db 1506 EAEKIVDESTIOELDEHTTASKNNEGKLNKDSIIRRLSENIEQNLNMLEKSAVK 1565
 QY 220 ELADRLKKDLVSSRSKIKLT-----NTELDQAKLELSAOKDLOQSADEITSLRKSD 272
 Db 1566 RLSTERSEILQFNRSRLADLEHKQSVESELSGSKLKLASTTEELQALAEKNERSLTTRML 1625
 QY 273 DPEGNLEPASATNETVSRVLFESAPAEVEMNPRLHQPPEFDEIDLNTTFDVPNPPPTQSG 332
 Db 1626 DLQNVQVDSLNIKDSLSE-----DLRLRLSLE--DSVASTL 1658
 QY 333 SOHCLPKLCLERARSPMNVLKKVH-----KVSKEPSQSLSGQRCVGLDEELAG 384
 Db 1659 QKECKIKSNTE-----SLQVLTYSQARNALEDEVSRSVDKIRRRDRC-----EHLTG 1709
 RESULT 16
 MYH9_HUMAN STANDARD; PRT; 1960 AA.
 AC P35579; 060805;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Myosin heavy chain, nonmuscle type A (Cellular myosin heavy chain,
 DE type A) (Nonmuscle myosin heavy chain-A) (NMHC-A).
 GN MYH9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
 RA Clamp M., Smink L.J., Alnsough R., Almeida J.P., Babbeay A.,
 RA Baggaley C., Bailey J., Barlow K., Bates K.N., Beasley O., Bird C.P.,
 RA Blakey S., Bridgeman A.M., Buck D., Burgess J., Burrill W.D.,
 RA Burton J., Carter C., Carter N.P., Chen Y., Clark G., Clegg S.M.,
 RA Cobley V., Cole C.G., Collier R.E., Connor R.E., Conroy D., Corby N.,
 RA Coville G.J., Cox A.V., Davis J., Dawson E., Dhand P.D., Dockree C.,
 RA Dowsorth S.J., Durbin R.M., Ellington A., Evans K.L., Fey J.M.,
 RA Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E.,
 RA Grafham D., Griffiths M.N., Hall C., Hall R., Hall-Tamlyn G.,
 RA Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J.,
 RA Kimberley A., King A., Laird G.K., Langford C.F., Leversha M.A.,
 RA Lloyd C., Lloyd D.M., Martyn I.D., Mashreghi-Mohammadi M.,
 RA Matthews L., Mccann O.T., Mcclay J., McLaren S., Mcmurray A.,
 RA Milne S.A., Mortimore B.J., Odell C.N., Pavitt R., Pearce A.V.,
 RA Pearson D., Phillimore B.J., Phillips S.H., Plumb R.W., Ramsay H.,
 RA Ramsey Y., Rogers L., Ross M.T., Scott C.E., Sehra H.K., Stuce C.D.,
 RA Smalley S., Smith M.L., Soderlund C., Spragon L., Stewart C.A.,
 RA Sulston J.E., Swann R.M., Vaudin M., Wall M., Wallis J.M.,
 RA Whiteley M.N., Willey D., Williams L., Williams S., Williamson H.,
 RA Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R.,
 RA Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T.,
 RA Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N.,
 RA Mitsuyama S., Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S.,
 RA Do A., Do T., Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S.,
 RA Lai H., Lao H.T., Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E.,
 RA Nguyen T., Pan H., Phan S., Qi S., Qian Y., Ray L., Ren Q., Shauli S.,
 RA Sloan D., Song L., Wang Q., Wang Y., Wang Z., White J., Willingham D.,
 RA Wu H., Yao Z., Zhan W., Zhang G., Chisoe S., Murray J., Miller N.,
 RA Min P., Fulton R., Johnson D., Bems G., Bentley D., Bradshaw H.,
 RA Bourne S., Cordes M., Du Z., Fulton L., Goela D., Graves T.,
 RA Hawkins J., Hinds K., Kemp K., Latreille P., Layman D., Ozersty P.,
 RA Rohlfing T., Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K.,
 RA Nelson J., Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,

RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tjahjuna Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 RN [2]
 RP SEQUENCE OF 1-1337 FROM N.A.
 RX MEDLINE=92003925; PubMed=1912569;
 RA Toothaker L.E., Gonzalez D.A., Tung N., Lemons R.S., Le Beau M.M.,
 RA Arnaud M.A., Clayton L.K., Tenen D.G.;
 RT "Cellular myosin heavy chain in human leukocytes: isolation of 5'
 RT cDNA clones, characterization of the protein, chromosomal
 RT localization, and upregulation during myeloid differentiation.";
 RL Blood 78:1826-1833(1991).
 RN [3]
 RP SEQUENCE OF 1-715 FROM N.A.
 RX MEDLINE=91316803; PubMed=1860190;
 RA Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K.,
 RA Gdula D., Adelstein R.S., Weir L.;
 RT "Human nonmuscle myosin heavy chains are encoded by two genes located
 RT on different chromosomes.";
 RL Circ. Res. 69:530-539(1991).
 RN [4]
 RP SEQUENCE OF 714-1960 FROM N.A.
 RX MEDLINE=90138958; PubMed=1967836;
 RA Saez C.G., Myers J.C., Shows T.B., Leinwand L.A.;
 RT "Human nonmuscle myosin heavy chain mRNA: generation of diversity
 RT through alternative polyadenylation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:1164-1168(1990).
 RN [5]
 RP VARIANT DFNAL7 HIS-705.
 RX MEDLINE=20489856; PubMed=11023810;
 RA Lalwani A.K., Goldstein J.A., Kelley M.J., Luxford W., Castelein C.M.,
 RA Mhatre A.N.;
 RT "Human nonsyndromic hereditary deafness DFNAL7 is due to a mutation in
 RT nonmuscle myosin MYH9.";
 RL Am. J. Hum. Genet. 67:1121-1128(2000).
 RN [6]
 RP VARIANTS MHA/FTNS/SBS K-93; C-702; C-1165; H-1424 AND K-1841.
 RX MEDLINE=20428192; PubMed=10973259;
 RA Serl M., Cusano M., Gangarossa S., Caridi G., Bordo D., Lo Nigro C.,
 RA Gliggeri G.M., Ravazzolo R., Savino M., Del Vecchio M., d'Apollito M.,
 RA Iolascon A., Zeltante L.L., Savoia A., Balducci C.L., Norris P.,
 RA Magrini U., Belletti S., Heath K.E., Babcock M., Gluckman M.J.,
 RA Aliprandis E., Bizaro N., Desnick R.J., Martignetti J.A.;
 RT "Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and
 RT Sebastian syndromes.";
 RL Nat. Genet. 26:103-105(2000).
 RN [7]
 RP VARIANTS MHA ILE-1155 AND IYS-1841.
 RX MEDLINE=20428193; PubMed=10973260;
 RA Kelley M.J., Javien W., Ortel T.L., Korczak J.F.;
 RT "Mutation of MYH9, encoding non-muscle myosin heavy chain A, in
 RT May-Hegglin anomaly.";
 RL Nat. Genet. 26:106-108(2000).
 CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
 CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
 CC CAPPING.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
 CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
 CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF MAY-HEGGLIN ANOMALY
 CC (MHA), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
 CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUCOCYTE INCLUSIONS.
 CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF FECHTNER SYNDROME
 CC (FTNS), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
 CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUCOCYTE INCLUSIONS.
 CC WITH ADDITIONAL ALPORT-LIKE CLINICAL FEATURES OF SENSORINEURAL
 CC DEAFNESS, CATARACTS AND NEPHRITIS.
 CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF SEBASTIAN SYNDROME

(SBS), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.
CC -1- DISEASE DEFECTS IN MYH9 ARE THE CAUSE OF AN AUTOSOMAL DOMINANT FORM OF NONCYTROMEGALOCYTIC DEAFNESS (DFNA17) WHICH IS CHARACTERIZED BY PROGRESSIVE HEARING IMPAIRMENT AND COCHLEOSACULAR DEGENERATION.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z82215; CAB05105.1; -
CC EMBL; M81105; AAA59888.1; -
CC EMBL; M69180; AAA61765.1; -
CC EMBL; M31013; AAA36349.1; -
CC HSSP; P08799; ILVK.
CC MIM; 160775; -
CC MIM; 153640; -
CC MIM; 155100; -
CC MIM; 603622; -
CC MIM; 605249; -
CC InterPro: IPR000048; IQ.
CC InterPro: IPR004009; Myosin_N.
CC InterPro: IPR002928; Myosin_tail.
CC InterPro: IPR002017; Spectrin.
CC InterPro: IPR001609; myosin_head.
CC Pfam: PF006612; IQ; 1.
CC Pfam: PF00063; myosin_head; 1.
CC Pfam: PF02736; Myosin_N; 1.
CC Pfam: PF01576; Myosin_tail; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC ProDom; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MYSC; 1.
CC PROSITE; PSS0096; IQ; 1.
CC Myosin: ATP-binding; Calmodulin-binding; Actin-binding;
CC Coiled coil; Alkylation; Multigene family; Disease mutation;
CC Deafness.
CC -----
CC DOMAIN 1 778
CC DOMAIN 779 808
CC DOMAIN 837 1926
CC NP_BIND 174 181
CC DOMAIN 654 676
CC MOD_RES 694 694
CC MOD_RES 704 704
CC VARIANT 93 93
CC VARIANT 702 702
CC VARIANT 705 705
CC VARIANT 1155 1155
CC VARIANT 1165 1165
CC VARIANT 1424 1424
CC VARIANT 1841 1841
CC CONFLICT 53 55
CC CONFLICT 660 660
CC CONFLICT 869 869
CC CONFLICT 931 931
CC CONFLICT 1240 1241
CC CONFLICT 1350 1350
CC CONFLICT 1764 1764
CC CONFLICT 1771 1771
CC -----
CC IO MYOSIN HEAD-LIKE.
CC COILED COIL (POTENTIAL).
CC ATP (POTENTIAL).
CC ACTIN-BINDING.
CC ALKYLATION (SH-1) (POTENTIAL).
CC ALKYLATION (SH-2) (POTENTIAL).
CC N -> K (IN MHR).
CC /FTId-VAR_010791.
CC R -> C (IN FTNS).
CC /FTId-VAR_010792.
CC R -> H (IN DFNA17).
CC /FTId-VAR_010793.
CC T -> I (IN MHR).
CC /FTId-VAR_010794.
CC R -> C (IN SBS).
CC /FTId-VAR_010795.
CC D -> H (IN FTNS).
CC /FTId-VAR_010796.
CC E -> K (IN MHR).
CC /FTId-VAR_010797.
CC EAI -> RGH (IN REF. 3).
CC T -> S (IN REF. 3).
CC T -> M (IN REF. 4).
CC C -> Y (IN REF. 4).
CC KG -> GR (IN REF. 4).
CC E -> EE (IN REF. 2).
CC T -> A (IN REF. 2).
CC S -> G (IN REF. 2).

SO SEQUENCE 1960 AA; 226531 MW; 588F84BB8C106E6F CRC64;
Query Match 7.48; Score 177; DB 1; Length 1960;
Best Local Similarity 19.08; Pred. No. 0.11;
Matches 107; Conservative 114; Mismatches 181; Indels 160; Gaps 21;

QY 48 QCRVQGGKTTIINFPLAOEENV-----LDAEF-LKNEIDSVKAOISQK 93
DB 1073 ELKQQLAKKE--EELQALANAVEEBAQKNALKKIRELSQISELQEDLESEASRKA 1130
QY 94 DREKRDSQAIIDTLRDLTLEERNATVESIQNALNKAEMLSYTK-----OMKFELEQ 144
DB 1131 EKORLDGEELTEALTELEDLTDSTAQOELRSKREDEVNLTKTLEBEAKTHAQIQEM 1190
QY 145 RQ-----DETKAREEHRKCKM-----KME-----QIELLOS-----QSEV 180
DB 1191 RQKHSQAVEELAEQLEQKRVKANLEKAKOTLENERGELANEVVLLQGGKDSHKRKRV 1250
QY 181 EEMIRDMGV-----GOSAVEQLAVYCVSLKREYMLKEARKATGELADRLKKDLVSSRSKL 236
DB 1251 EAQQLQYKNEBERKRTTELADKVTKLQVELDVTGLISQSDSKSKLTDFSALESQ 1310
QY 237 K-----TINTELDQAKLELSAQKDLQSDAQ-----EITSRLR 268
DB 1311 QDTQELLQENRQKLSLSTKLQYEDENKSPREQLLEEEEEKHNLEKQIATLHAQVADMK 1370
QY 269 KKSDDPPGNLEPASATN-----ETVSRLVFESPA---PYEMNPRRHQPPFDEIDL 318
DB 1371 KKMEDSVGCTETAEVEVRKLOKDLGLSQREHEKVAAYDKLEKTKTRLQOELDLDDL 1430
QY 319 -----TTPDVNPPPTQSGSH-----CLP 338
DB 1431 HQRSACNLEKKQKFDQLAAEKTISKYAEEDRAEAREKETALSLARLEAME 1490
QY 339 KKLCLERA---RSPMONVL-----KVKHVSKEPSQSLSGQRC---VGLDEELAG 384
DB 1491 QKAELERLNKFRTMEDLMSSKDVGSVHELKSRALQOEKMKTOLELEDELQA 1550
QY 385 AFPLFIRNAVIGOKOPRTTAESSRSDVVRITGDLGGRKFPQPDITIRRV-PYKS 443
DB 1551 TEDAKRLREV-----NIQAMKQFERD-----LOGRDESEKKKOLVRQVREMA 1596
QY 444 KAKSKQKVRKTVSSASOPKLD 465
DB 1597 ELDERKQKRSMAVAKKLEMD 1618

RESULT 17
RA50_THEVO STANDARD; PRT; 895 AA.
ID RA50_THEVO
AC PS8302;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN RAD50 OR TVG0235331.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
OC Thermoplasma.
RN NCBI_TaxID=50339;
RP SEQUENCE FROM N.A.
RC STRAIN=GS51 / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Nunoshima T., Yamamoto Y., Aramaki H., Kanehori K., Kawamoto T.,
RT "Archaeal adaptation to higher temperatures revealed by genomic
sequence of Thermoplasma volcanium";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
rad50/mre11 complex possesses single-strand endonuclease activity

```

CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP000991; BAB59370.1; -.
CC InterPro: IPR000439; ABC_transportr.
CC InterPro: IPR000875; Cecropin.
CC InterPro: IPR001238; RecF.
CC InterPro: IPR003395; SMC_N.
CC Pfam: PF00470; RecF. 1.
CC Pfam: PF02463; SMC_N. 1.
CC DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
CC NP_BIND 30
CC FT DOMAIN 171 737 Coiled coil (POTENTIAL).
CC S0 SEQUENCE 895 AA; 103223 MW; 4A47DA9287E82D3A CRC64;

Query Match 7.4%; Score 176.5; DB 1; Length 895;
Best Local Similarity 20.5%; Pred. No. 0.049;
Matches 75; Conservative 77; Mismatches 135; Indels 79; Gaps 12;

OY 51 IOVGKKTINKLPFDLAO-----EEENVLDA-----EFKNELDVSVAQ 89
DB 122 LNIKKDVFLLSIYRQGEMLAVSKDPAERKKIIDEILINRLAGYILLLEVIDDLTAN 181
OY 90 LSQDKREKRSQAIITDLRLTEERNATVESLQNALNKAEMLCSTLKKOMKFLPQRODET 149
DB 182 VSDYDYLNELQSKT-----NEIDNNKKQIELEESKRLRIEPEIKALEEINIKENKKD-- 235
OY 150 KQAEAEHRLCKMKMTQETILL---QSQSEVEEMIRDMGVGQSAVEQLAVCVSLKK 206
DB 236 -HLNEELHRLNAQLETKYKMEELAESQSKASIEMEVYKLPSTIEELKRLENNAAVYKR 294
OY 207 -----EVENLKEARKATGELADRLK-----KDIYSSRSKL---KTLNTELDQA 246
DB 295 NEIEFYINLKDLQSLSEIILGKLSLSKYDEAHRKLEDDLSFSEPLEKKRKREDLDKL 354
OY 247 KLELRSAQKDLQSAQDEITSLRKSDDPGNLEPASATNETVSRVLESPAPVEMMPRL 306
DB 355 RSLKEDEDNYQSAVRNIENIKWIEKEIERMSAFISILKIQETLP---ETINSR- 410
OY 307 HQPFGEIDINTFTDVTPTQTSGSHCLPKKLCLEFRANSPMONVLKKHKYKSPESQ 366
DB 411 -----RAEINSSLMQIEG-----KIASLNASIDAMRS-----HKMEVEENA 446
OY 367 LSLGGO 372
DB 447 AMLSGR 452

RESULT 18
MYSN ACACA STANDARD; PRT; 1509 AA.
AC POS659;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin II heavy chain, non muscle.
OS Acanthamoeba castellanii (Ameoba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
ON NCBI_taxid=5755;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=87308395; PubMed=3040773;
RA Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;
RT "complete nucleotide sequence and deduced polypeptide sequence of a
RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT hinge in the rodlike tail."
RL J. Cell Biol. 105:913-925(1987).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES.
CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC -1- THE REGULATORY PHOSPHORYLATION SITES RESIDE.
CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMAIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y00624; CA68663.1; -.
CC PIR: A27224; A27224.
CC HSP: P08799; IMND.
CC InterPro: IPR000048; IQ.
CC InterPro: IPR004009; Myosin_N.
CC InterPro: IPR001609; myosin_head.
CC Pfam: PF00612; IQ; 2.
CC Pfam: PF00663; myosin_head; 1.
CC Pfam: PF02736; Myosin_N; 1.
CC PRINTS: PR00193; MYOSINHEAVY.
CC PRODOM: PD000355; myosin_head; 1.
CC SMART: SM00015; IQ; 1.
CC SMART: SM00242; MYSC; 1.
CC DR PROSITE: PSS0096; IQ; 1.
CC KW Myosin; coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
CC Methylation; Alkylation; Phosphorylation; Multigene family.
CC MYOSIN HEAD-LIKE.
CC FT DOMAIN 1 789
CC FT DOMAIN 790 819
CC FT DOMAIN 848 1509
CC FT DOMAIN 848 1226
CC FT DOMAIN 1227 1252
CC FT DOMAIN 1253 1509
CC FT DOMAIN 1253 1482
CC FT DOMAIN 1483 1509
CC FT NP_BIND 182 189
CC FT DOMAIN 660 682
CC FT DOMAIN 766 780
CC FT MOD_RES 133 133
CC FT MOD_RES 700 700
CC FT MOD_RES 1489 1489
CC FT MOD_RES 1494 1494
CC FT MOD_RES 1499 1499
CC S0 SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 7.4%; Score 176.5; DB 1; Length 1509;
Best Local Similarity 24.5%; Pred. No. 0.086;
Matches 81; Conservative 63; Mismatches 113; Indels 73; Gaps 15;

OY 66 LAQEEENVLDAEFLKNELDVSVAQKLSQDKREKRSQAIITDLRLTEERN-----TYESL 121
DB 888 LAEE-----DADKLEKDLAALKRLIDEGEKAD-----LEEDNALLOKKVAGL 931

```


| | | | |
|------------|------------|--|---------------|
| QY | 122 | ONALNKABMLCSTLTKKMKAKLEORODETKARPEAHRLKCMKMTYEQEOLLQSORSEVE | 181 |
| Db | 932 | EEELQEEETSA5NDLELEOKRRLLEAKESKELNALSIEEER--- | 988 |
| QY | 182 | EMIDRMCGOSAVEQJLAVYCVSLKKEYENL-KEARKATGELAD--RLKDKLVSSRSKAKT | 238 |
| Db | 989 | DKYDEEAAND-----SLKKKEEDLSRLRETRKALADAENISITL--- | 10355 |
| QY | 239 | -----LNTLELDO--AKLELRSAOKDLQASADQETTSLLRKSDDPPGNILEPASFATNEV | 288 |
| Db | 1036 | TERGADYVRNELDVTATKLOLEKTKKSLT---EEELAOQTRALDEEKSQGEAA5SKAKOL | 10922 |
| QY | 289 | SRVFE5PAPYEMNNPRLHQPPFGDELDLNTFPPDVTPPTQSG5GQHCLKPKLCLERARS | 348 |
| Db | 1093 | GQOLEDARSEVDSLKSLT5---AAEK5LKTAKQDNRLDEO-----LEDEBT | 1136 |
| QY | 349 | PMQNVLRKVH-----KVSKPESQL-SLGGOR | 373 |
| Db | 1137 | VRANVDKOKALEKLTLELDQVTAIDGOK | 1166 |
| RESULT | 19 | | |
| MYSS_RABIT | | STANDARD; | PRT; 1084 AA. |
| ID | MYSS_RABIT | | |

AC P02562; (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, last sequence update)
 DT 01-JAN-1990 (Rel. 13, last annotation update)
 DT 15-JUL-1998 (Rel. 36, last annotation update)
 DE Myosin heavy chain, skeletal muscle (fragments).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxID=9986;

RN [1]
RP SEQUENCE OF 1-258.
RA Capony J.-P., Elzinga M.;
RT "The amino acid sequence of A 34,000 dalton fragment from S-2 of
RL myosin.",
BIophys. J. 33:148a-148a(1981).

RN [2]
RP SEQUENCE OF 259-428.
RX MEDLINE=65131142; PubMed=3972832;
RA Lu R.C., Wong A.;
RT "The amino acid sequence and stability predictions of the hinge
region in myosin subfragment 2.";
J. Biol. Chem. 260:3456-3461(1985).

```

RN      [3]
RP      SEQUENCE OF 409-1084 FROM N.A.
RX      MEDLINE=87304245; PubMed=3305014;

```

RA Maeda K., Szczakiel G., Wittinghofer A.;
RT "Characterization of cDNA coding for the complete light meromyosin
RT portion of a rabbit fast skeletal muscle myosin heavy chain.",
RL Eur. J. Biochem. 167:97-102(1987).

CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUPERFILIAR LOCATION: Thick filaments of the myofibrils

-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC
CC MYOMYOSIN (LMM) AND 1 HEAVY MYOMYOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC
CC -1- CRYSTALLINE AND NONCRYSTALLINE OR AMORPHOUS AND CHAIN-EXTENDED

CC WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE
CC CONSERVED.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```
CC -----
DR EMBL; X05958; CAA29391.1; -.
DR PTR: A02985; A02985.
```

DR PIR; A05280; A05280.
DR PIR; S00084; S00084.
DR PIR; S00084; S00084.

DR Pfam; PF01576; Myosin_tail; 1.
KM Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;

| Accession | Protein | Length (aa) | PI | MW (kDa) | Domain |
|-----------|--------------------------------|-------------|----|----------|--------------------------------|
| KW010000 | ATP-binding; Multigene family. | 1 | 1 | 25.8 | ATP-binding; Multigene family. |
| FT010000 | NON-TER | 1 | 1 | 25.8 | NON-TER |
| FT010000 | DOMAIN | 1 | 1 | 25.8 | DOMAIN |

| | | | |
|----------|------|------|-------------------------------|
| NON_CONS | 258 | 259 | RODLIKE TAIL (S2 AND LMM DOM) |
| DOMAIN | <259 | 1084 | |
| FT | 1084 | | |
| FT | | | |
| DOMAIN | | | |

| | | | |
|----|-----|------|--------|
| FT | 455 | 1084 | COILED |
| FT | 405 | 405 | L -> V |
| FT | 408 | 408 | V -> L |

| | | | | |
|----|---------|-----|-----|--------|
| FT | VARIANT | 421 | 421 | E -> D |
| FT | VARIANT | 423 | 423 | S -> G |
| FT | VARIANT | 426 | 426 | K -> P |

| FI | VARIANT | 420 | 420 | N | N |
|----|----------|------|-----|------------------|-------|
| SQ | SEQUENCE | 1084 | AA; | 125488 | MM; |
| | | | | 229CFD69A6E1F7F0 | CRC64 |

```
Query Match      7.48; Score 176; DB 1; Length 1084
Best Local Similarity 20.78; Pred. No. 0.064;
```

| Match | Conservative | Mismatches | Indels | Gaps |
|-------|--------------|------------|--------|------|
| 82 | 85 | 163 | 66 | 13 |

[illegible]

OY 101 QAI--IDTLRDTLEERNATVESLQNALINKAEMLCSTLKOKMFLEORODETKOAREBAH

Db 455 QAF7QOIEGKRLQLEETTKAKSALAHALQSSRRDCDLUREQYEEEEQAKELORAMSKANN

```

Db      515 SEVSOMTTCETDAIORTE-ELEEAKKLAORLOD---AEENVEAVNSKASLEKTORIL
          |||:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy      158 -----RLCKRMKTMCIELLLOSOREVEEMIRDMGSGOSAVEQLVAVCSLKRKEYENL

```

QY 212 KEARKATGELADRLKKDLVSSRSKLTUNNE-----LDOAKLELRSAOKDLOSADO 267

```

Db      571 Q--NEADLMIDVERSNATCARMDKKQRNEDKVLAEKHKYEETQELASQKEERSLST 638

```

QY 263 EITSLRKKSDDBPNLEPASATNETVSRVFEESPAPVEMNPNRLHQ-PPFGDEIDLNTTF 321

QY 322 DVNTPRTQTSQHCLPKKLCLERARSPMÖNVLKVKHKVSKPESÖLSLGGRCVGLDEE 381

```

Db      684 -----QKSELQALAEAGSLHEEGKILRIQLLELVQKSEIDRKINKEDEE 731

```

```

OY      382 LAGAPLEFIRNAV-LGQKQPNRTAESRSSSTDVRI 41
      :      || : : : : || || | : ||
Db      732 T----DQKPNHIVVFESMOSTIDAEITSPNDALPT 76

```

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

| | | |
|-----------|-----------|------------------------|
| RESULT | 20 | |
| MYP_CAEEL | | |
| ID | MYP_CAEEL | STANDARD: PRT: 882 AA. |

AC P10567;
DT 01-JUL-1989 (Rel. 11, Created)
BT 01-JUL-1989 (Rel. 11, Last Accessed: 06/01/2001)

| | |
|----|---|
| DE | Paramyosin. |
| DT | 15-JUL-1998 (Rel. 36, Last annotation update) |
| DI | 01-JUL-1989 (Rel. 11, Last sequence update) |

GN UNC-15.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita.

OC Rhaditidae; Pelodierinae; Caenorhabditis.
NCBI_TaxID=6239; OX

RN [1]

```

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99329026; PubMed-2754728;
RA Kagawa H., Gengyo K., McLachlan A.D., Brenner S., Karn J.;
RT "Paramyosin gene (unc-15) of Caenorhabditis elegans. Molecular
RT cloning, nucleotide sequence and models for thick filament
RT structure."
RT J. Mol. Biol. 207:311-333(1989).
RN [2]
RP ERRATUM.
RA Kagawa H., Gengyo K., McLachlan A.D., Brenner S., Karn J.;
RN J. Mol. Biol. 211:665-665(1990).
RN [3]
RP PHOSPHORYLATION.
RX MEDLINE-99329036; PubMed-2754733;
RA Schrieffer L.A., Waterson R.H.;
RT "Phosphorylation of the N-terminal region of Caenorhabditis elegans
RT paramyosin."
RL J. Mol. Biol. 207:451-454(1989).
CC -1- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
CC MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: FOR MOST OF ITS LENGTH, PARAMYOSIN APPEARS TO FORM AN
CC ALPHA-HELICAL COILED COIL, AND SHOWS THE HEPTAD REPEAT OF
CC CHARGED AMINO ACIDS CHARACTERISTIC OF MYOSIN HEAVY CHAINS.
CC HOWEVER, PARAMYOSIN DIFFERS FROM MYOSIN IN HAVING NON-HELICAL
CC EXTENSIONS AT BOTH TERMINI AND AN ADDITIONAL "SKIP" RESIDUE THAT
CC INTERRUPTS THE 28-RESIDUE REPEAT. THE DISTRIBUTION OF CHARGED
CC RESIDUES IS ALSO DIFFERENT FROM MYOSIN HEAVY CHAINS.
CC -1- PPM: PHOSPHORYLATED ON SERINE RESIDUES IN THE NON-ALPHA-HELICAL
CC N-TERMINAL REGION.
CC -1- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X08068; CAA30857.1; -.
DR PIR: S04027; S04027.
DR HSSP: P80220; IDIP.
DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PF01576; Myosin_tail; 1.
KW Coiled coil; Muscle protein; Thick filament; Myosin; Phosphorylation.
FT DOMAIN 1 47 NONHELICAL REGION (POTENTIAL).
FT DOMAIN 48 871 COILED COIL (POTENTIAL).
FT DOMAIN 872 882 NONHELICAL REGION (POTENTIAL).
FT DISULFID 143 143 INTERCHAIN (POTENTIAL).
FT DISULFID 632 632 INTERCHAIN (POTENTIAL).
SQ SEQUENCE 882 AA; 101949 MW; 894C4BFF4A670C5F CRC64;

```

```

Query Match 7.3%; Score 175; DB 1; Length 882;
Best Local Similarity 21.8%; Pred. No. 0.057;
Matches 96; Conservative 65; Mismatches 136; Indels 144; Gaps 17;

```

```

QY 195 EQLAVCVSLKKEYENLKEARK-----ATGELAD--RIKKDLYSSRSKLTINTLTD 244
| : : : : : | : : : : : | : : : : : | : : : : : |
DB 623 AALSABEELCEATLADNIAIRAKQAEVDFLEANGSLISLNNLSTIKKLE---TELS 679
QY 245 QAKLELSAQDKLOLSDAQ-----ETSLRKSSDDPPGULP 280
| : : : : : | : : : : : | : : : : : | : : : : : |
DB 680 TAQADLDEVTKELHAADERANRALADARAVEQLHEQDEHSMKIDALRKSLSEGVKQLQ- 738
QY 281 ASATNETHVSRLVFESPAVEMNPLHQPPGDEIDINTFTDVNTPPTQTSNGSHCLPKP 340
| : : : : : | : : : : : | : : : : : | : : : : : |
DB 739 -VQIEAEPAALLCGRKYATLKEFRIR-----DETALDEET----- 774
QY 341 LCLERRSPMQNVLYKKVYKSPESQSLGQRCVGEDELAEAPFLIRNAVLCQKOP 400
| : : : : : | : : : : : | : : : : : | : : : : : |
DB 775 ---RRKHEQNALRKDKDRKEVQL-----VDEE-----HKNFMAQDTA 812
QY 401 NRTT-----AESRSST 411
| : : : : : | : : : : : | : : : : : | : : : : : |
DB 813 DRLTEKINIKRQLAESSEVT 833

```

RESULT 21

```

MYH9_RAT STANDARD; PRT; 1961 AA.
ID AC Q62812;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle type A (cellular myosin heavy chain,
DE type A) (Nonmuscle myosin heavy chain-A).
GN MYH9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Choi O.H., Park C., Itoh K., Adelstein R.S., Beaven M.A.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPING.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES.
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 TO DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U31463; AAA74950.1; -.
DR HSSP: P10587; IBRI.
DR InterPro: IPR000048; IO.
DR InterPro: IPR000409; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IO; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IO; 1.
DR SMART: SM00242; MYSC; 1.

```

DR PROSITE; PS50096; IQ; 1.
 KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
 KM Colled coil; Alkylation; Multigene family.
 FT DOMAIN 1 778 MYOSIN HEAD-LIKE.
 FT DOMAIN 779 808
 FT DOMAIN 841 1927 COILED COIL (POTENTIAL).
 FT NP_BIND 174 181 ATP (POTENTIAL).
 FT DOMAIN 654 676 ACTIN-BINDING.
 FT MOD_RES 694 694 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
 SO SEQUENCE 1961 AA; 226336 MW; 989876D9681FBI9E CRC64;

Query Match 7.3%; Score 174.5; DB 1; Length 1961;
 Best Local Similarity 19.4%; Pred. No. 0.14;
 Matches 109; Conservative 112; Mismatches 180; Indels 161; Gaps 21;

QY 48 QCRIOVGKTTINKLFPLDAGEENV-----LDNEF--LKNELDSVKAQLSQK 93
 DB 1073 ELKMOIAKE--EELQALALAVEEAAQKNMALKKIRELETOISELDSEERACRNKA 1130
 QY 94 DREKRSQAIIIDLTLDTEENATVESIQNALNKAEMLCTLTK-----QMKPLEQ 144
 DB 1131 EKOKRDGLCEELALKTLELDLSTAAQOELRSKREQEVSLIKTLEDEANTHEAQIDEM 1190
 QY 145 RO-----DETQAREEAHRLCKM---KTME-----QIELLSQ-----QRSEV 180
 DB 1191 ROKNSQAVEELAEQLEQKRRKATLEKAKOTLENRGELANVAKLQCKDSEIKRRKV 1250
 QY 181 EEMTRDMGV---GQSAVEQLAVVCSLKEKEYENKEARKATGELADRIKLDVSSRKL 236
 DB 1251 EAQIQELQVKEFSEGERVTELDADKYSKLQVELDSVTGLNDSKSKITLDFALSQSL 1310
 QY 237 K-----TUNTELDOAKTELRSOKDLSADQ-----ITSL 267
 DB 1311 QDTQELQEEKROKISLSTKLQMEDEKNSFREQLDEEEKKNLEKQIATLHAQVTDW 1370
 QY 268 RKKSDDPPGNLEPASATN-----ETVSRLVFESPA---PYEMNPRTHOPPFGEIDL 317
 DB 1371 KKKMEDVGCLSETAEAKRRLQKDELSQRLSEKVAAYDLKTKTQLQELDLVDL 1430
 QY 318 N-----TTTPVNTPPPTQSGQH-----CL 337
 DB 1431 DHORQSVNLEKKOKKFFQLAEKTTISAKYAEERDRAEAAREKETALSLARALEBAM 1490
 QY 338 PKKICLERA---RSPMONVL-----KVVHAKSPRESQSLSGQRC---VGELEDELA 383
 DB 1491 EOKAELEFLNKQFTMEDLSSKDDVGKSVHELEKSNRALEDOVEKKTQLELEDELQ 1550
 QY 384 GAPPLEFIRNAVLAGQKOPARTTAESRSSTDVVRIGFDGLGRTKFTIOPRDTTIIIRPVVKS 443
 DB 1551 ATEDAKLRLVEY-----NLQAKMAQFERD-----LQGDDEQSEKKKQLVNRQVR-EM 1595
 QY 444 KAKSKOKVRIITVSSASQPKD 465
 DB 1596 EAELDEKORKSIAAARKKL 1617

RESULT 22
 XCPE_XENLA STANDARD; PRT; 1203 AA.
 ID XCPE_XENLA
 AC P50533;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Chromosome assembly protein XCAP-E.
 GN XCAP-E.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 ON NCBI_TaxID=8355;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-95042742; PubMed-7954811;
 RX Hirono T., Mitchison T.J.:
 RT "A heterodimeric coiled-coil protein required for mitotic chromosome
 condensation in vitro".
 RL Cell 79:449-458(1994).
 CC -I- FUNCTION: REQUIRED FOR BOTH ASSEMBLY AND STRUCTURAL MAINTENANCE OF
 MITOTIC CHROMOSOMES.
 CC -I- SUBUNIT: ASSOCIATES WITH XCAP-C PROBABLY AS HETERODIMER.
 CC -I- SUBCELLULAR LOCATION: NUCLEAR. DURING CHROMOSOME ASSEMBLY IN
 MITOTIC EXTRACTS, XCAP-C/E WAS RECRUITED TO THE CHROMATIN AND
 FORMED A DISCRETE INTERNAL STRUCTURE WITHIN ASSEMBLED CHROMOSOMES.
 CC -I- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS
 FLANKED BY PUTATIVE GLOBAL REGIONS AT THE N- AND C-TERMINUS.
 CC -I- SIMILARITY: BELONGS TO THE SMC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; U13674; AAA64680.1; .
 DR HSSP; P07751; 1A33.
 DR InterPro; IPR003405; SMC_C.
 DR InterPro; IPR003395; SMC_N.
 DR Pfam; PF02483; SMC_C; 1.
 DR Pfam; PF02463; SMC_N; 1.
 KW Mitosis; ATP-binding; Coiled coil; Nuclear protein.
 FT NP_BIND 32 39 ATP (POTENTIAL).
 FT DOMAIN 172 356 COILED COIL (POTENTIAL).
 FT DOMAIN 397 513 COILED COIL (POTENTIAL).
 FT DOMAIN 670 1032 COILED COIL (POTENTIAL).
 FT DOMAIN 1095 1121 ALA/ASP-RICH (DA-BOX).
 SO SEQUENCE 1203 AA; 136341 MW; 04323DD0027DF309 CRC64;

Query Match 7.3%; Score 173.5; DB 1; Length 1203;
 Best Local Similarity 21.9%; Pred. No. 0.095;
 Matches 73; Conservative 69; Mismatches 135; Indels 57; Gaps 11;

QY 77 EFLNELDSVAQOISQKREKRSQAIIIDLTLDTEENATVESIQNALNKAEMLCTLTK 136
 DB 721 EMKSEALDLOTKLQOSSHKQOE--LDSLKQITESEELTKTKVEQKAAEFKYLE 778
 QY 137 KOMKFLF-QRODETRQARE-----EAHRLCKMKTMEQIELLSQSRSEVEEMTRD 186
 DB 779 HKMKAEAEERRELEKQAKIDTAKKRAASNNKKMEKQOEVDAVL-----ELEELKRE 833
 QY 187 MGQSAVEQLAVVCSLKEKEYENKEARKATGELADRIKLDVSSRKLTKLTINTELDOA 246
 DB 834 Q-----TTYKQOIEVDENAKAYQDQADSMASEVSKNKAARKADELAKQ 879
 QY 247 KLELRSAOKDLSADOETISLRKKSDDPPGNLEPASATNTEVSRLVFESPAV-EMNPR- 304
 DB 880 KEIIMGHDKETIKTSSEAGKIRENNNDOLKIKIELEHITSKHKDSADAAKAVKMLNDY 939
 QY 305 --RIHOPPFGEIDLTNTTTPVNTPPPTQSGQ--HCLPRKL-----CLEARAR 347
 DB 940 EMIASEKHLFG--QANTAYDFKTNPNKQAGORLHKLOEKKEKIGRNVNMRAMNLTQAE 996
 QY 348 SPMONVLKVVAVKSPRESQSLSGQRCVGELEDE 381
 DB 997 ERYNDLMKRRRIVENDSKI---LTITIEIDOK 1026

RESULT 23
 MYH8_HUMAN STANDARD; PRT; 1937 AA.
 ID MYH8_HUMAN
 AC P13535; Q14910;
 DT 01-JAN-1990 (Rel. 13, Created)

DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Myosin heavy chain, skeletal muscle, perinatal (MYHC-perinatal).
GN MYH8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90323631; PubMed=2373371;
RA Karsch-Mizrachi I., Feghali R., Shows T.B. Jr., Leinwand L.A.;
RT "Generation of a full-length human perinatal myosin heavy-chain-
RT encoding cDNA.";
RL Gene 89:289-294(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95324556; PubMed=7601129;
RA Jullien E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,
RA Stedman H.H., Rubinstein N.A.;
RT "Characterization of a human perinatal myosin heavy-chain
RT transcript.";
RL Eur. J. Biochem. 230:1001-1006(1995).
RN [3]
RP SEQUENCE OF 502-1937 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
RT myosin heavy chains.";
RL Eur. J. Biochem. 189:55-65(1990).
RN [4]
RP SEQUENCE OF 860-1937 FROM N.A.
RX MEDLINE=89234168; PubMed=2715179;
RA Feghali R., Leinwand L.A.;
RT "Molecular genetic characterization of a developmentally regulated
RT human perinatal myosin heavy chain.";
RL J Cell Biol. 108:1791-1797(1989).
RN [5]
RP SEQUENCE OF 1-46 FROM N.A.
RA Esser K., Tidhar A., Myszkowski M.;
RT "Isolation and characterization of the human perinatal MHC promoter.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC AKTYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROWOSIN (LMM) AND 1 HEAVY MEROWOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M36769; AAC17185.1; -;
DR EMBL; Z38133; CAA86293.1; -;

DR EMBL; X51592; CAA35941.1; -;
DR EMBL; M35250; AAA36346.1; -;
DR EMBL; AF067143; AAC21557.1; -;
DR PIR; A30220; A30220.
DR HSSP; P08799; ILVK.
DR MIM; 160741; -;
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR01609; myosin_head.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PSS0096; IQ; 1.
KW Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family;
KW Calmodulin-binding.
FT DOMAIN 1 780 MYOSIN HEAD-LIKE.
FT DOMAIN 781 813 IQ.
FT DOMAIN 842 1937 COILED COIL (POTENTIAL).
FT NP_BIND 181 188 ATP.
FT DOMAIN 658 680 ACTIN-BINDING.
FT DOMAIN 760 774 ACTIN-BINDING.
FT MOD_RES 132 132 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 698 698 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 708 708 ALKYLATION (SH-2) (POTENTIAL).
FT MOD_RES 15 15 A -> R (IN REF. 2).
FT CONFLICT 970 970 M -> N (IN REF. 1 AND 4).
FT CONFLICT 1072 1072 N -> H (IN REF. 3).
FT CONFLICT 1247 1247 MC -> DGG (IN REF. 3).
FT CONFLICT 1251 1252 E -> G (IN REF. 1 AND 4).
FT CONFLICT 1261 1261 K -> Q (IN REF. 1 AND 4).
FT CONFLICT 1297 1297 KY -> NT (IN REF. 3).
FT CONFLICT 1377 1378 EN -> AH (IN REF. 1 AND 4).
FT CONFLICT 1504 1505 E -> D (IN REF. 1 AND 4).
FT CONFLICT 1847 1847 D -> H (IN REF. 2).
FT CONFLICT 1914 1914
SQ SEQUENCE 1937 AA; 222762 MW; A3EE2D151792E9E8 CRC64;

Query Match 7.2%; Score 172.5; DB 1; Length 1937;
Best Local Similarity 19.8%; Pred. No. 0.18;
Matches 77; Conservative 79; Mismatches 167; Indels 65; Gaps 9;

QY 55 KKTIIINKLFFDLAOEEENVDAEFLKNELDSVKAOLDSQDKREKRDQAIIIDTLRDTLEER 114
DB 1269 QORLINDL---TAQRARLQGTAGEYSKQDLDEKDALVSLSKASQAOIEELKHOLEEE 1325
QY 115 NATVESLONALNKAEMLCSTLKQMKFLEQRODETKQAREEAHRIKCKIKTMEQIEILLQ 174
DB 1326 TKAKNALALHLOSSRHDDLLREQYEEBQGAELQRLSKANSEAVQMRKYETFAIOR 1385
QY 175 SORSEVEE---MIRDKVGSAVEQOLAVYCVSLKKEVENK-----EARKATGE 220
DB 1386 TE--ELEBAKKKLQORLOEAEBHEAVAVAKCASLEKTKORLONEVEDMLDVERSNACA 1443
QY 221 LADRLKK--DLVSSRSKKTINTELDOAKLELRSAOKDLOSAODEITSLRKSDPPGNL 278
DB 1444 ALDKKQRFDFVLSWKOK-----YEELQALELSQKESRSLSTELRKVKKNVYESLDL 1498
QY 279 EPASATNETVSRLVFEESPAPVEMNPRLOPPFQDEIDLNTTFDVNTPPTQTSQSQHCLP 338
DB 1499 ETLRRENKNLQOEISDLTEQIABEGKQIHE-----LEKIKQYGEKCE 1542
QY 339 KKLCLERARSPQNVKLKVHVKYSKRESOLSLGGORCVGELDEELAGAPPLFIRNAVLGOK 398
DB 1543 IOALEEAEASILEHEGKILRIQLELQNVKSEVVRKIAEKDEE-----IDL 1589

OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Miller N.;
 RU Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: WEAK, TO MYOSINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U41545; AKK39135.1; -
 DR WormRep; C02F12.7; CE03901.
 DR InterPro; IPR003727; DUF149.
 DR Pfam; PF02575; DUF149; 1.
 KW Hypothetical protein; Coiled coil.
 FT DOMAIN 121 779 COILED COIL (POTENTIAL).
 FT DOMAIN 805 1061 COILED COIL (POTENTIAL).
 SQ SEQUENCE 1130 AA; 131485 MW; B0FD2EF3D99FB09 CRC64;

Query Match 7.1%; Score 171; DB 1; Length 1130;
 Best Local Similarity 20.0%; Pred. No. 0.12;
 Matches 92; Conservative 88; Mismatches 189; Indels 92; Gaps 15;
 OY 60 NKLFFDLAQR-----ENVLDAEF--LNKELDSVKAQLSQDKRERDQAIIIDT 106
 DB 612 NRLHENTQKLEAHKETHETVRLAEIDQFKSAFENEQDYGEKSAKIRELEAQKKT 671
 OY 107 LRDLEENAVESIQNALNKAEMLCSTLKQMKFLERQDETQCARERARLCKM--- 163
 DB 672 LLSERKVKHVAENLEAFTSDKNLLELESKNKNIEHLKQELQNLNKKISTKTEKQSE 731
 OY 164 --KTMQTEILLQSQRSEVEEM---IRDMGVQSAVEQALVAYCVSLKVEYNLKEARKAT 218
 DB 732 LKTLAQLEIDNSSKSDQIEKLHLRVNDM-----LDMGKITKDELYKKNIEKITISKT 785
 OY 219 GELADRLKKDLVSRSKIKITNTLELDQAKLELSAQDKLQASADQI----- 264
 DB 786 AQL---LESNTVESETKLASYTEEREK---EIOSFQTIISKEDNEVLTKARINLEETCL 839
 OY 265 -----TSLRKSPDPGNGLEPASAT---NETYSRLVFESPPAPYEMNPLHQ-----P 309
 DB 840 KEREVELTGKTKLDDMTQOLNEETVYVLFNNSIOEKIDKEATTNENKRLKSRNEIA 899
 OY 310 PFGDEIDLNTTFDVTPTQTSQHCLPKKLCLEFARSPQNVLYKVKHVV-----SKP 363
 DB 900 KLHEEMVQKQKNEKRNQSKLQELMFEKEQLEAKESHIIEAVEQYFQADKEKXW 959
 OY 364 ESQSLSGQRCVGLDELFLACAPFLIRNAVVGQKQPKRTTAESRSSYDVVRIGFDGLGG 423
 DB 960 KEQI-----EDLENA--LQKRNELIQLODRQTESSEPTKRMSTISHGV 1005
 OY 424 RTKFI-OPRDTTIRPVYKSKAKSKOKVRKITYSASQPK 463
 DB 1006 FQNFVSQMKD-----KREASEKRTKREAEKKAKEK 1037

RESULT 26
 MYH4_HUMAN STANDARD; PRT: 1939 AA.
 AC Q91623;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, skeletal muscle, fetal (Myosin heavy chain I1b)
 DE (MYHC-I1b).
 GN MYH4.

OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=99318869; PubMed=10388558;
 RA Weiss A., Schiaffino S., Leinwand L.A.;
 RT "Comparative sequence analysis of the complete human sarcomeric myosin
 RT heavy chain family: implications for functional diversity.";
 RL J. Mol. Biol. 290:61-75(1999).
 CC -1- SUBUNIT: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF111783; AAD29949.1; -
 DR HSSP; P08799; IMMD.
 DR MIM; 160742; -
 DR InterPro; IPR000048; IQ.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR002928; Myosin_tail.
 DR InterPro; IPR001609; myosin_head.
 DR Pfam; PF00612; IQ; 2.
 DR Pfam; PF00063; myosin_head; 1.
 DR Pfam; PF02736; Myosin_N; 1.
 DR Pfam; PF01576; Myosin_tail; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS50096; IQ; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
 FT DOMAIN 785 814 IO.
 FT DOMAIN 843 1939 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 SQ SEQUENCE 1939 AA; 223012 MW; 40B1ADID77A47DE CRC64;

Query Match 7.1%; Score 170.5; DB 1; Length 1939;
 Best Local Similarity 20.5%; Pred. No. 0.22;
 Matches 78; Conservative 83; Mismatches 169; Indels 51; Gaps 11;
 OY 55 KTIINKLFFDLAQRSEENVLDAEFLKNEIDSVKAQLSQDKRERDQAIIIDTLRDTLEER 114
 DB 1270 QORLINE---SAQKARLHETSGERSQLDEKDAVNSQLSGKQAFQOIEELKRLQLEE 1326
 OY 115 NATVESLONALNKAEMLCSTLKQMKFLERQDETQCARERARLCKMKTMEQTEILL-- 172


```

FT REPEAT 2648 2685 PLECTIN 15.
FT REPEAT 2724 2761 PLECTIN 16.
FT REPEAT 2762 2799 PLECTIN 17.
FT DOMAIN 2824 2847 6 x 4 AA TANDEM REPEATS OF G-S-R-[SR].
FT DOMAIN 1 584 INTERACTS WITH PLAKOPHILIN 1 AND JUNCTION
FT PLAKOLOBIN.
FT MOD_RES 2849 2849 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT VARSPIC 1195 1794 MISSING (IN ISOFORM DP11).
FT CONFLICT 905 905 R -> A (IN REF. 2).
FT CONFLICT 1120 1120 D -> R (IN REF. 3).
FT CONFLICT 2687 2688 SV -> RL (IN REF. 2).
SQ SEQUENCE 2871 AA; 331771 MW; 5639B5B7CD4690B7 CRC64;

```

```

Query Match 7.1%; Score 170.5; DB 1; Length 2871;
Best Local Similarity 18.0%; Pred. No. 0.34; Mismatches 210; Indels 103; Gaps 15;
Matches 93; Conservative 110;

```

```

QY 35 IOWETAPSRTPCQRIQVQSKTIINKLFPDLAOEENVLDLAEF-----78
DB 1514 LQKANSATETINKLQVQDELTRL-RIDYERVQERTVQKDDITRFQNSLKELOKOK 1572
QY 79 LKNEIDSVKQLOSKDKREKDSQAIITLRLDTEERNATVESLQNALNKAEMLC-----ST 134
DB 1573 VEELNLKRTKASDSCKRRKLEEGMRSLKEQAIKITNLQOLEQASIVKRRSDD 1632
QY 135 LKKQMKLEORODETKOAREAHRLCKMKMTQEOIELLLOSRS-----EVE 181
DB 1633 LRQRRVLDGHLRKQKQTEELRLSSVEALRR--OLLOQESVKQAHNLNEHFQKAYE 1690
QY 182 EMIRDMGVGSAVBOLAVYCVSLKKEYENLKE-----ARKATGELADRLKDLV 230
DB 1691 DKRSLSNESKLEIRLOSLFENLTKHEMLLEELRNLRLEYDRLRGSEADSKNNATIL 1750
QY 231 SSRKLTNTLDELQAKLELSAQKDLOSADQETTSLSKSSDPPGNLEPA-----SATN 285
DB 1751 ELRSQLOISNNRLELQGLINDLOREBNLEKOEIKFOKALFASNNRLOESKNOCTOVVQ 1810
QY 286 ETVSRIV-----FESPAPVEMNRLHOPPGDELDLNT-----FDVNTPTQ 329
DB 1811 ERELILKIKVLEODKRLQLEDELNRAKSTLAEIRVKORLECEKQOIONDLNOMKTO 1870
QY 330 TSGSOHCLPKKLCLELRASP-----MONVLKKVHK-----VSKPESOLS 368
DB 1871 YSRKEEAI-RKIESERESEKSNLSSEIRLQAEIKRIEERCRKLEDDSTRETQGLE 1929
QY 369 LGGRCVGELEDELAGAPFLFIRNAVIGQKOPNRTAESRSSTVVRIGFDLGGRKEI 428
DB 1930 TERSRYOREIDK-----LRORPYGSHRETQT--ECEWTVDTSKLVFEDLKKKYAM 1978
QY 429 QPRDTTIIRPVVSKSKAKSKOKVRIKTPVSSASQPKL 464
DB 1979 QLYECQLDKTTLDKLLKGGKSV--EYVASIOLPFL 2012

```

```

RESULT 28
TMFL_HUMAN STANDARD; PRT; 1093 AA.
AC P82094;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TATP element modulatory factor (TMF).
GN TMF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. AND FUNCTION.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=93028466; PubMed=1409643;
RA Garcia J.A., Ou S.-H.I., Wu F., Lusis A.J., Sparkes R.S., Gaynor R.B.;

```

```

RT "Cloning and chromosomal mapping of a human immunodeficiency virus 1
RT 'TATP' element modulatory factor."
RL Proc. Natl. Acad. Sci. U.S.A. 89:9372-9376(1992).
CC -FUNCTION: THIS PROTEIN BINDS THE HIV-1 TATP ELEMENT AND INHIBITS
CC TRANSCRIPTIONAL ACTIVATION BY THE TATP-BINDING PROTEIN (TBP).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L01042; AAD54608.1; -.
DR MTM; 601126; -.
KW Transcription regulation; DNA-binding; Repressor; Coiled coil.
FT DOMAIN 439 922 COILED COIL (POTENTIAL).
FT DOMAIN 984 1092 COILED COIL (POTENTIAL).
SQ SEQUENCE 1093 AA; 123170 MW; 26133B85F4677BE CRC64;

```

```

Query Match 7.1%; Score 170; DB 1; Length 1093;
Best Local Similarity 20.2%; Pred. No. 0.13;
Matches 107; Conservative 79; Mismatches 203; Indels 140; Gaps 17;

```

```

QY 66 LAQEEENVLDLAEFLKNEIDSVKQLOSKDKREKDSQAIITLRLT---ERNATVSLQ 122
DB 441 LSEKEDCVKEVFEFLNEKLEKREQLLSLSEKALLEAFNLDKEMRVEESSISLSK 500
QY 123 N-----ALNKAEMLC---STLKQMK-----FLEGRQDSTKQARE 154
DB 501 DEFQRTAEAEKKYQLOACKERDAKKEIKNKEELATRLNNSETFADLKEDEIDRLME 560
QY 155 EAHRLCKMKMTQEOIELLOSQSEVEEMIRDMGVGSAVBOLAVYCVSLKKEYENLKEA 214
DB 561 EGEKLSQQLHNSNIKKLRADKKNENMVAKL---NKVYLEEELQHLKQVLDGKKEV 617
QY 215 RKATGELADRLKDKLVSSRSKLTNLDELQAKLELSAQKDLOSADQETTSLSKSSDPP 274
DB 618 EKQRENIKTKLNSVVERQEKDGLRQVMDLEKKNNSIQALDASVAKETLDLHKANAAK 677
QY 275 PGNLEPASATNETSRVLFESPAPVEMNRLHOP-----PFGD----- 313
DB 678 DSEQGEALSLSEMKAK--EELSLALEKAQEFARQOQETIATQVDDLRALQRTQQAARK 735
QY 314 -----ET-----DLNTTFDVNTPP-----TQTSQSOHCLPKKL-- 341
DB 736 EDYLRRHEIGELQRLQEAENRNOELSSQSVSTPRPLRLQIENLQATLGSSWEKLEKN 795
QY 342 -----CLEKARPMQNVLKVKYHKVSKPESQSLGQ---RCVGELEDELAG 384
DB 796 LSDRLGESQTLAAVAVERATTELLANKIQMSMESQNLRLQENSRFQALESE-- 852
QY 385 AFPLFIRNAVIGQKOPN-----RTTASRSSTDVVRIGFDLGGRKEI 429
DB 853 -----KNRLCKLEDENNRKYVLEENLKDEYVRLTEERKRTKTLINSOLE--MERMAVEQ 904
QY 430 PRDTTIIRPVVSKSKAKSKOKV-----RIKTVSSASQPKLDT-FIQ 470
DB 905 ERKKAIFQETIKERKRPFSVSTPTMSRSSISGVDMAGLQNSFLSQ 953

```

```

RESULT 29
HIPL_HUMAN STANDARD; PRT; 995 AA.
AC O00291; 000328;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Huntingtin interacting protein 1 (HIP-1) (Fragment).
GN HIP1.
OS Homo sapiens (Human).

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hug A.H.M., Nichol K., Osborne L., Scherer S.W., Squitieri F.,
 RA Hayden M.R.;
 RT "Genomic organization of the human HIP1 gene and its exclusion as a
 RT candidate gene in a family diagnosed with Huntington disease without
 RT CAG expansion";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 82-995 FROM N.A.
 RX MEDLINE=97285121; PubMed=9140394;
 RA Kaichman M.A., Koide H.B., McCutcheon K., Graham R.K., Nichol K.,
 RA Nishiyama K., Kazemi-Esfarjani P., Lynn F.C., Wellington C.,
 RA Metzler M., Goldberg Y.P., Kanazawa I., Gelfz R.D., Hayden M.R.;
 RT "HIP1, a human homologue of S. cerevisiae Slap2, interacts with
 RT membrane-associated huntingtin in the brain";
 RL Nat. Genet. 16:44-53(1997).
 RN [3]
 RP SEQUENCE OF 82-159 FROM N.A.
 RA Bradshaw H., Hinds K., Harrison M.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 203-602 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97227296; PubMed=9147654;
 RA Wanker E.E., Rovira C., Scherzinger E., Hasenbank R., Welter S.,
 RA Tait D., Colicelli J., Lehner H.;
 RT "HIP-1: a huntingtin interacting protein isolated by the yeast two-
 RT hybrid system";
 RL Hum. Mol. Genet. 6:487-495(1997).
 RN [5]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=20515263; PubMed=11063258;
 RA Chopra V.S., Metzler M., Rasper D.M., Engqvist-Goldstein A.E.Y.,
 RA Singaraja R., Gan L., Fichter K.M., McCutcheon K., Drubin D.,
 RA Nicholson D.W., Hayden M.R.;
 RT "HIP2 is a non-proapoptotic member of a gene family including HIP1,
 RT an interacting protein with huntingtin";
 RL Mamm. Genome 11:1006-1013(2000).
 CC -1- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE CELL FILAMENT
 CC NETWORKS.
 CC -1- SUBUNIT: BINDS HUNTINGTIN. THIS INTERACTION IS RESTRICTED TO THE
 CC BRAIN.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. MEMBRANE-ASSOCIATED PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HIP1-1 AND HIP1-2; ARE PRODUCED
 CC BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED WITH THE HIGHEST LEVEL
 CC IN BRAIN.
 CC -1- MISCELLANEOUS: THE AFFINITY OF THE HUNTINGTIN PROTEIN-HIP1
 CC INTERACTION IS INVERSELY CORRELATED TO THE LENGTH OF THE
 CC POLYGLUTAMINE TRACT ADDED TO THE HUNTINGTIN PROTEIN IN HUNTINGTON
 CC DISEASE.
 CC -1- SIMILARITY: BELONGS TO THE SLA2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF052268; AAC33564.1; JOINED.
 DR EMBL: AF052269; AAC33564.1; JOINED.
 DR EMBL: AF052270; AAC33564.1; JOINED.
 DR EMBL: AF052271; AAC33564.1; JOINED.
 DR EMBL: AF052272; AAC33564.1; JOINED.
 DR EMBL: AF052273; AAC33564.1; JOINED.
 DR EMBL: AF052274; AAC33564.1; JOINED.
 DR EMBL: AF052275; AAC33564.1; JOINED.
 DR EMBL: AF052276; AAC33564.1; JOINED.
 DR EMBL: AF052277; AAC33564.1; JOINED.
 DR EMBL: AF052278; AAC33564.1; JOINED.
 DR EMBL: AF052279; AAC33564.1; JOINED.
 DR EMBL: AF052280; AAC33564.1; JOINED.
 DR EMBL: AF052281; AAC33564.1; JOINED.
 DR EMBL: AF052282; AAC33564.1; JOINED.
 DR EMBL: AF052283; AAC33564.1; JOINED.
 DR EMBL: AF052284; AAC33564.1; JOINED.
 DR EMBL: AF052285; AAC33564.1; JOINED.
 DR EMBL: AF052286; AAC33564.1; JOINED.
 DR EMBL: AF052287; AAC33564.1; JOINED.
 DR EMBL: AF052288; AAC33564.1; JOINED.
 DR EMBL: AF052289; AAC33564.1; JOINED.
 DR EMBL: AF052290; AAC33564.1; JOINED.
 DR EMBL: AF052291; AAC33564.1; JOINED.
 DR EMBL: AF052292; AAC33564.1; JOINED.
 DR EMBL: AF052293; AAC33564.1; JOINED.
 DR EMBL: AF052294; AAC33564.1; JOINED.
 DR EMBL: AF052295; AAC33564.1; JOINED.
 DR EMBL: AF052296; AAC33564.1; JOINED.
 DR EMBL: AF052297; AAC33564.1; JOINED.
 DR EMBL: AF052298; AAC33564.1; JOINED.
 DR EMBL: AF052299; AAC33564.1; JOINED.
 DR EMBL: AF052300; AAC33564.1; JOINED.
 DR EMBL: AF052301; AAC33564.1; JOINED.
 DR EMBL: AF052302; AAC33564.1; JOINED.
 DR EMBL: AF052303; AAC33564.1; JOINED.
 DR EMBL: AF052304; AAC33564.1; JOINED.
 DR EMBL: AF052305; AAC33564.1; JOINED.
 DR EMBL: AF052306; AAC33564.1; JOINED.
 DR EMBL: AF052307; AAC33564.1; JOINED.
 DR EMBL: AF052308; AAC33564.1; JOINED.
 DR EMBL: AF052309; AAC33564.1; JOINED.
 DR EMBL: AF052310; AAC33564.1; JOINED.
 DR EMBL: AF052311; AAC33564.1; JOINED.
 DR EMBL: AF052312; AAC33564.1; JOINED.
 DR EMBL: AF052313; AAC33564.1; JOINED.
 DR EMBL: AF052314; AAC33564.1; JOINED.
 DR EMBL: AF052315; AAC33564.1; JOINED.
 DR EMBL: AF052316; AAC33564.1; JOINED.
 DR EMBL: AF052317; AAC33564.1; JOINED.
 DR EMBL: AF052318; AAC33564.1; JOINED.
 DR EMBL: AF052319; AAC33564.1; JOINED.
 DR EMBL: AF052320; AAC33564.1; JOINED.
 DR EMBL: AF052321; AAC33564.1; JOINED.
 DR EMBL: AF052322; AAC33564.1; JOINED.
 DR EMBL: AF052323; AAC33564.1; JOINED.
 DR EMBL: AF052324; AAC33564.1; JOINED.
 DR EMBL: AF052325; AAC33564.1; JOINED.
 DR EMBL: AF052326; AAC33564.1; JOINED.
 DR EMBL: AF052327; AAC33564.1; JOINED.
 DR EMBL: AF052328; AAC33564.1; JOINED.
 DR EMBL: AF052329; AAC33564.1; JOINED.
 DR EMBL: AF052330; AAC33564.1; JOINED.
 DR EMBL: AF052331; AAC33564.1; JOINED.
 DR EMBL: AF052332; AAC33564.1; JOINED.
 DR EMBL: AF052333; AAC33564.1; JOINED.
 DR EMBL: AF052334; AAC33564.1; JOINED.
 DR EMBL: AF052335; AAC33564.1; JOINED.
 DR EMBL: AF052336; AAC33564.1; JOINED.
 DR EMBL: AF052337; AAC33564.1; JOINED.
 DR EMBL: AF052338; AAC33564.1; JOINED.
 DR EMBL: AF052339; AAC33564.1; JOINED.
 DR EMBL: AF052340; AAC33564.1; JOINED.
 DR EMBL: AF052341; AAC33564.1; JOINED.
 DR EMBL: AF052342; AAC33564.1; JOINED.
 DR EMBL: AF052343; AAC33564.1; JOINED.
 DR EMBL: AF052344; AAC33564.1; JOINED.
 DR EMBL: AF052345; AAC33564.1; JOINED.
 DR EMBL: AF052346; AAC33564.1; JOINED.
 DR EMBL: AF052347; AAC33564.1; JOINED.
 DR EMBL: AF052348; AAC33564.1; JOINED.
 DR EMBL: AF052349; AAC33564.1; JOINED.
 DR EMBL: AF052350; AAC33564.1; JOINED.
 DR EMBL: AF052351; AAC33564.1; JOINED.
 DR EMBL: AF052352; AAC33564.1; JOINED.
 DR EMBL: AF052353; AAC33564.1; JOINED.
 DR EMBL: AF052354; AAC33564.1; JOINED.
 DR EMBL: AF052355; AAC33564.1; JOINED.
 DR EMBL: AF052356; AAC33564.1; JOINED.
 DR EMBL: AF052357; AAC33564.1; JOINED.
 DR EMBL: AF052358; AAC33564.1; JOINED.
 DR EMBL: AF052359; AAC33564.1; JOINED.
 DR EMBL: AF052360; AAC33564.1; JOINED.
 DR EMBL: AF052361; AAC33564.1; JOINED.
 DR EMBL: AF052362; AAC33564.1; JOINED.
 DR EMBL: AF052363; AAC33564.1; JOINED.
 DR EMBL: AF052364; AAC33564.1; JOINED.
 DR EMBL: AF052365; AAC33564.1; JOINED.
 DR EMBL: AF052366; AAC33564.1; JOINED.
 DR EMBL: AF052367; AAC33564.1; JOINED.
 DR EMBL: AF052368; AAC33564.1; JOINED.
 DR EMBL: AF052369; AAC33564.1; JOINED.
 DR EMBL: AF052370; AAC33564.1; JOINED.
 DR EMBL: AF052371; AAC33564.1; JOINED.
 DR EMBL: AF052372; AAC33564.1; JOINED.
 DR EMBL: AF052373; AAC33564.1; JOINED.
 DR EMBL: AF052374; AAC33564.1; JOINED.
 DR EMBL: AF052375; AAC33564.1; JOINED.
 DR EMBL: AF052376; AAC33564.1; JOINED.
 DR EMBL: AF052377; AAC33564.1; JOINED.
 DR EMBL: AF052378; AAC33564.1; JOINED.
 DR EMBL: AF052379; AAC33564.1; JOINED.
 DR EMBL: AF052380; AAC33564.1; JOINED.
 DR EMBL: AF052381; AAC33564.1; JOINED.
 DR EMBL: AF052382; AAC33564.1; JOINED.
 DR EMBL: AF052383; AAC33564.1; JOINED.
 DR EMBL: AF052384; AAC33564.1; JOINED.
 DR EMBL: AF052385; AAC33564.1; JOINED.
 DR EMBL: AF052386; AAC33564.1; JOINED.
 DR EMBL: AF052387; AAC33564.1; JOINED.
 DR EMBL: AF052388; AAC33564.1; JOINED.
 DR EMBL: AF052389; AAC33564.1; JOINED.
 DR EMBL: AF052390; AAC33564.1; JOINED.
 DR EMBL: AF052391; AAC33564.1; JOINED.
 DR EMBL: AF052392; AAC33564.1; JOINED.
 DR EMBL: AF052393; AAC33564.1; JOINED.
 DR EMBL: AF052394; AAC33564.1; JOINED.
 DR EMBL: AF052395; AAC33564.1; JOINED.
 DR EMBL: AF052396; AAC33564.1; JOINED.
 DR EMBL: AF052397; AAC33564.1; JOINED.
 DR EMBL: AF052398; AAC33564.1; JOINED.
 DR EMBL: AF052399; AAC33564.1; JOINED.
 DR EMBL: AF052400; AAC33564.1; JOINED.
 DR EMBL: AF052401; AAC33564.1; JOINED.
 DR EMBL: AF052402; AAC33564.1; JOINED.
 DR EMBL: AF052403; AAC33564.1; JOINED.
 DR EMBL: AF052404; AAC33564.1; JOINED.
 DR EMBL: AF052405; AAC33564.1; JOINED.
 DR EMBL: AF052406; AAC33564.1; JOINED.
 DR EMBL: AF052407; AAC33564.1; JOINED.
 DR EMBL: AF052408; AAC33564.1; JOINED.
 DR EMBL: AF052409; AAC33564.1; JOINED.
 DR EMBL: AF052410; AAC33564.1; JOINED.
 DR EMBL: AF052411; AAC33564.1; JOINED.
 DR EMBL: AF052412; AAC33564.1; JOINED.
 DR EMBL: AF052413; AAC33564.1; JOINED.
 DR EMBL: AF052414; AAC33564.1; JOINED.
 DR EMBL: AF052415; AAC33564.1; JOINED.
 DR EMBL: AF052416; AAC33564.1; JOINED.
 DR EMBL: AF052417; AAC33564.1; JOINED.
 DR EMBL: AF052418; AAC33564.1; JOINED.
 DR EMBL: AF052419; AAC33564.1; JOINED.
 DR EMBL: AF052420; AAC33564.1; JOINED.
 DR EMBL: AF052421; AAC33564.1; JOINED.
 DR EMBL: AF052422; AAC33564.1; JOINED.
 DR EMBL: AF052423; AAC33564.1; JOINED.
 DR EMBL: AF052424; AAC33564.1; JOINED.
 DR EMBL: AF052425; AAC33564.1; JOINED.
 DR EMBL: AF052426; AAC33564.1; JOINED.
 DR EMBL: AF052427; AAC33564.1; JOINED.
 DR EMBL: AF052428; AAC33564.1; JOINED.
 DR EMBL: AF052429; AAC33564.1; JOINED.
 DR EMBL: AF052430; AAC33564.1; JOINED.
 DR EMBL: AF052431; AAC33564.1; JOINED.
 DR EMBL: AF052432; AAC33564.1; JOINED.
 DR EMBL: AF052433; AAC33564.1; JOINED.
 DR EMBL: AF052434; AAC33564.1; JOINED.
 DR EMBL: AF052435; AAC33564.1; JOINED.
 DR EMBL: AF052436; AAC33564.1; JOINED.
 DR EMBL: AF052437; AAC33564.1; JOINED.
 DR EMBL: AF052438; AAC33564.1; JOINED.
 DR EMBL: AF052439; AAC33564.1; JOINED.
 DR EMBL: AF052440; AAC33564.1; JOINED.
 DR EMBL: AF052441; AAC33564.1; JOINED.
 DR EMBL: AF052442; AAC33564.1; JOINED.
 DR EMBL: AF052443; AAC33564.1; JOINED.
 DR EMBL: AF052444; AAC33564.1; JOINED.
 DR EMBL: AF052445; AAC33564.1; JOINED.
 DR EMBL: AF052446; AAC33564.1; JOINED.
 DR EMBL: AF052447; AAC33564.1; JOINED.
 DR EMBL: AF052448; AAC33564.1; JOINED.
 DR EMBL: AF052449; AAC33564.1; JOINED.
 DR EMBL: AF052450; AAC33564.1; JOINED.
 DR EMBL: AF052451; AAC33564.1; JOINED.
 DR EMBL: AF052452; AAC33564.1; JOINED.
 DR EMBL: AF052453; AAC33564.1; JOINED.
 DR EMBL: AF052454; AAC33564.1; JOINED.
 DR EMBL: AF052455; AAC33564.1; JOINED.
 DR EMBL: AF052456; AAC33564.1; JOINED.
 DR EMBL: AF052457; AAC33564.1; JOINED.
 DR EMBL: AF052458; AAC33564.1; JOINED.
 DR EMBL: AF052459; AAC33564.1; JOINED.
 DR EMBL: AF052460; AAC33564.1; JOINED.
 DR EMBL: AF052461; AAC33564.1; JOINED.
 DR EMBL: AF052462; AAC33564.1; JOINED.
 DR EMBL: AF052463; AAC33564.1; JOINED.
 DR EMBL: AF052464; AAC33564.1; JOINED.
 DR EMBL: AF052465; AAC33564.1; JOINED.
 DR EMBL: AF052466; AAC33564.1; JOINED.
 DR EMBL: AF052467; AAC33564.1; JOINED.
 DR EMBL: AF052468; AAC33564.1; JOINED.
 DR EMBL: AF052469; AAC33564.1; JOINED.
 DR EMBL: AF052470; AAC33564.1; JOINED.
 DR EMBL: AF052471; AAC33564.1; JOINED.
 DR EMBL: AF052472; AAC33564.1; JOINED.
 DR EMBL: AF052473; AAC33564.1; JOINED.
 DR EMBL: AF052474; AAC33564.1; JOINED.
 DR EMBL: AF052475; AAC33564.1; JOINED.
 DR EMBL: AF052476; AAC33564.1; JOINED.
 DR EMBL: AF052477; AAC33564.1; JOINED.
 DR EMBL: AF052478; AAC33564.1; JOINED.
 DR EMBL: AF052479; AAC33564.1; JOINED.
 DR EMBL: AF052480; AAC33564.1; JOINED.
 DR EMBL: AF052481; AAC33564.1; JOINED.
 DR EMBL: AF052482; AAC33564.1; JOINED.
 DR EMBL: AF052483; AAC33564.1; JOINED.
 DR EMBL: AF052484; AAC33564.1; JOINED.
 DR EMBL: AF052485; AAC33564.1; JOINED.
 DR EMBL: AF052486; AAC33564.1; JOINED.
 DR EMBL: AF052487; AAC33564.1; JOINED.
 DR EMBL: AF052488; AAC33564.1; JOINED.
 DR EMBL: AF052489; AAC33564.1; JOINED.
 DR EMBL: AF052490; AAC33564.1; JOINED.
 DR EMBL: AF052491; AAC33564.1; JOINED.
 DR EMBL: AF052492; AAC33564.1; JOINED.
 DR EMBL: AF052493; AAC33564.1; JOINED.
 DR EMBL: AF052494; AAC33564.1; JOINED.
 DR EMBL: AF052495; AAC33564.1; JOINED.
 DR EMBL: AF052496; AAC33564.1; JOINED.
 DR EMBL: AF052497; AAC33564.1; JOINED.
 DR EMBL: AF052498; AAC33564.1; JOINED.
 DR EMBL: AF052499; AAC33564.1; JOINED.
 DR EMBL: AF052500; AAC33564.1; JOINED.
 DR EMBL: AF052501; AAC33564.1; JOINED.
 DR EMBL: AF052502; AAC33564.1; JOINED.
 DR EMBL: AF052503; AAC33564.1; JOINED.
 DR EMBL: AF052504; AAC33564.1; JOINED.
 DR EMBL: AF052505; AAC33564.1; JOINED.
 DR EMBL: AF052506; AAC33564.1; JOINED.
 DR EMBL: AF052507; AAC33564.1; JOINED.
 DR EMBL: AF052508; AAC33564.1; JOINED.
 DR EMBL: AF052509; AAC33564.1; JOINED.
 DR EMBL: AF052510; AAC33564.1; JOINED.
 DR EMBL: AF052511; AAC33564.1; JOINED.
 DR EMBL: AF052512; AAC33564.1; JOINED.
 DR EMBL: AF052513; AAC33564.1; JOINED.
 DR EMBL: AF052514; AAC33564.1; JOINED.
 DR EMBL: AF052515; AAC33564.1; JOINED.
 DR EMBL: AF052516; AAC33564.1; JOINED.
 DR EMBL: AF052517; AAC33564.1; JOINED.
 DR EMBL: AF052518; AAC33564.1; JOINED.
 DR EMBL: AF052519; AAC33564.1; JOINED.
 DR EMBL: AF052520; AAC33564.1; JOINED.
 DR EMBL: AF052521; AAC33564.1; JOINED.
 DR EMBL: AF052522; AAC33564.1; JOINED.
 DR EMBL: AF052523; AAC33564.1; JOINED.
 DR EMBL: AF052524; AAC33564.1; JOINED.
 DR EMBL: AF052525; AAC33564.1; JOINED.
 DR EMBL: AF052526; AAC33564.1; JOINED.
 DR EMBL: AF052527; AAC33564.1; JOINED.
 DR EMBL: AF052528; AAC33564.1; JOINED.
 DR EMBL: AF052529; AAC33564.1; JOINED.
 DR EMBL: AF052530; AAC33564.1; JOINED.
 DR EMBL: AF052531; AAC33564.1; JOINED.
 DR EMBL: AF052532; AAC33564.1; JOINED.
 DR EMBL: AF052533; AAC33564.1; JOINED.
 DR EMBL: AF052534; AAC33564.1; JOINED.
 DR EMBL: AF052535; AAC33564.1; JOINED.
 DR EMBL: AF052536; AAC33564.1; JOINED.
 DR EMBL: AF052537; AAC33564.1; JOINED.
 DR EMBL: AF052538; AAC33564.1; JOINED.
 DR EMBL: AF052539; AAC33564.1; JOINED.
 DR EMBL: AF052540; AAC33564.1; JOINED.
 DR EMBL: AF052541; AAC33564.1; JOINED.
 DR EMBL: AF052542; AAC33564.1; JOINED.
 DR EMBL: AF052543; AAC33564.1; JOINED.
 DR EMBL: AF052544; AAC33564.1; JOINED.
 DR EMBL: AF052545; AAC33564.1; JOINED.
 DR EMBL: AF052546; AAC33564.1; JOINED.
 DR EMBL: AF052547; AAC33564.1; JOINED.
 DR EMBL: AF052548; AAC33564.1; JOINED.
 DR EMBL: AF052549; AAC33564.1; JOINED.
 DR EMBL: AF052550; AAC33564.1; JOINED.
 DR EMBL: AF052551; AAC33564.1; JOINED.
 DR EMBL: AF052552; AAC33564.1; JOINED.
 DR EMBL: AF052553; AAC33564.1; JOINED.
 DR EMBL: AF052554; AAC33564.1; JOINED.
 DR EMBL: AF052555; AAC33564.1; JOINED.
 DR EMBL: AF052556; AAC33564.1; JOINED.
 DR EMBL: AF052557; AAC33564.1; JOINED.
 DR EMBL: AF052558; AAC33564.1; JOINED.
 DR EMBL: AF052559; AAC33564.1; JOINED.
 DR EMBL: AF052560; AAC33564.1; JOINED.
 DR EMBL: AF052561; AAC33564.1; JOINED.
 DR EMBL: AF052562; AAC33564.1; JOINED.
 DR EMBL: AF052563; AAC33564.1; JOINED.
 DR EMBL: AF052564; AAC33564.1; JOINED.
 DR EMBL: AF052565; AAC33564.1; JOINED.
 DR EMBL: AF052566; AAC33564.1; JOINED.
 DR EMBL: AF052567; AAC33564.1; JOINED.
 DR EMBL: AF052568; AAC33564.1; JOINED.
 DR EMBL: AF052569; AAC33564.1; JOINED.
 DR EMBL: AF052570; AAC33564.1; JOINED.
 DR EMBL: AF052571; AAC33564.1; JOINED.
 DR EMBL: AF052572; AAC33564.1; JOINED.
 DR EMBL: AF052573; AAC33564.1; JOINED.
 DR EMBL: AF052574; AAC33564.1; JOINED.
 DR EMBL: AF052575; AAC33564.1; JOINED.
 DR EMBL: AF052576; AAC33564.1; JOINED.
 DR EMBL: AF052577; AAC33564.1; JOINED.
 DR EMBL: AF052578; AAC33564.1; JOINED.
 DR EMBL: AF052579; AAC33564.1; JOINED.
 DR EMBL: AF052580; AAC33564.1; JOINED.
 DR EMBL: AF052581; AAC33564.1; JOINED.
 DR EMBL: AF052582; AAC33564.1; JOINED.
 DR EMBL: AF052583; AAC33564.1; JOINED.
 DR EMBL: AF052584; AAC33564.1; JOINED.
 DR EMBL: AF052585; AAC33564.1; JOINED.
 DR EMBL: AF052586; AAC33564.1; JOINED.
 DR EMBL: AF052587; AAC33564.1; JOINED.
 DR EMBL: AF052588; AAC33564.1; JOINED.
 DR EMBL: AF052589; AAC33564.1; JOINED.
 DR EMBL: AF052590; AAC33564.1; JOINED.
 DR EMBL: AF052591; AAC33564.1; JOINED.
 DR EMBL: AF052592; AAC33564.1; JOINED.
 DR EMBL: AF052593; AAC33564.1; JOINED.
 DR EMBL: AF052594; AAC33564.1; JOINED.
 DR EMBL: AF052595; AAC33564.1; JOINED.
 DR EMBL: AF052596; AAC33564.1; JOINED.
 DR EMBL: AF052597; AAC33564.1; JOINED.
 DR EMBL: AF052598; AAC33564.1; JOINED.
 DR EMBL: AF052599; AAC33564.1; JOINED.
 DR EMBL: AF052600; AAC33564.1; JOINED.
 DR EMBL: AF052601; AAC33564.1; JOINED.
 DR EMBL: AF052602; AAC33564.1; JOINED.
 DR EMBL: AF052603; AAC33564.1; JOINED.
 DR EMBL: AF052604; AAC33564.1; JOINED.
 DR EMBL: AF052605; AAC33564.1; JOINED.
 DR EMBL: AF052606; AAC33564.1; JOINED.
 DR EMBL: AF052607; AAC33564.1; JOINED.
 DR EMBL: AF052608; AAC33564.1; JOINED.
 DR EMBL: AF052609; AAC33564.1; JOINED.
 DR EMBL: AF052610; AAC33564.1; JOINED.
 DR EMBL: AF052611; AAC33564.1; JOINED.
 DR EMBL: AF052612; AAC33564.1; JOINED.
 DR EMBL: AF052613; AAC33564.1; JOINED.
 DR EMBL: AF052614; AAC33564.1; JOINED.
 DR EMBL: AF052615; AAC33564.1; JOINED.
 DR EMBL: AF052616; AAC33564.1; JOINED.
 DR EMBL: AF052617; AAC33564.1; JOINED.
 DR EMBL: AF052618; AAC33564.1; JOINED.
 DR EMBL: AF052619; AAC33564.1; JOINED.
 DR EMBL: AF052620; AAC33564.1; JOINED.
 DR EMBL: AF052621; AAC33564.1; JOINED.
 DR EMBL: AF052622; AAC33564.1; JOINED.
 DR EMBL: AF052623; AAC33564.1; JOINED.
 DR EMBL: AF052624; AAC33564.1; JOINED.
 DR EMBL: AF052625; AAC33564.1; JOINED.
 DR EMBL: AF052626; AAC33564.1; JOINED.
 DR EMBL: AF052627; AAC33564.1; JOINED.
 DR EMBL: AF052628; AAC33564.1; JOINED.
 DR EMBL: AF052629; AAC33564.1; JOINED.
 DR EMBL: AF052630; AAC33564.1; JOINED.
 DR EMBL: AF052631; AAC33564.1; JOINED.
 DR EMBL: AF052632; AAC33564.1; JOINED.
 DR EMBL: AF052633; AAC33564.1; JOINED.
 DR EMBL: AF052634; AAC33564.1; JOINED.
 DR EMBL: AF052635; AAC33564.1; JOINED.
 DR EMBL: AF052636; AAC33564.1; JOINED.
 DR EMBL: AF052637; AAC33564.1; JOINED.
 DR EMBL: AF052638; AAC33564.1; JOINED.
 DR EMBL: AF052639; AAC33564.1; JOINED.
 DR EMBL: AF052640; AAC33564.1; JOINED.
 DR EMBL: AF052641; AAC33564.1; JOINED.
 DR EMBL: AF052642; AAC33564.1; JOINED.
 DR EMBL: AF052643; AAC33564.1; JOINED.
 DR EMBL


```

CC -----
DR EMBL: D89992; BAA22069.1; -
DR EMBL: D50476; BAA09069.1; -
DR EMBL: D43700; BAA07802.1; -
DR HSSP: P08799; IMKD.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;
KW Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IQ.
FT DOMAIN 812 839 HINGE.
FT DOMAIN 840 1935 COILED COIL (POTENTIAL).
FT NP_BIND 178 185 ATP (POTENTIAL).
FT DOMAIN 659 681 ACTIN-BINDING.
FT DOMAIN 761 775 ACTIN-BINDING.
FT MOD_RES 129 129 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 699 699 ALKYLATION (SH-1).
FT MOD_RES 709 709 ALKYLATION (SH-2).
SQ SEQUENCE 1935 AA; 221599 MW; 9A1244B67D63C83B CRC64;

```

Query Match 7.18; Score 169; DB 1; Length 1935;
 Best Local Similarity 19.38; Pred. No. 0.26; Mismatches 173; Indels 138; Gaps 18;
 Matches 97; Conservative 90;

```

QY 31 LQCLWMEFAPSRPCPCRIQGVGKTIINKLFPLAQEEVLD-----AEFLNELDS 85
DB 1144 LEISERLEEGAGTAAIDENKKKEAEFOKKRRD---EESTLOHEKTAALAKREQDS 1200
QY 86 VKAQLSQK-----DREKDSQAIIDTLRDTLEERNATVESLONALKAKMELCST 134
DB 1201 V-AEIGEIDMLQRYKQKLEKSEYKMEIDL-----TSNBAVAKAKANLEKMCRT 1252
QY 135 LKKQKKFLEQRODETKQAREHNRKCKMKT-----MGQIEL----- 172
DB 1253 LEDQSEIKTKSDENVROLNPMNARARLQTENGEFSKOLEKEALVSQLTFRGQAYTQQ 1312
QY 173 LOSORSEVEEMIRDMGVGQSAVEOLAVVCSLKEKENLKEAR-----KATGELAD- 223
DB 1313 IEELEKRIHEEVKAKNALAHVQSAHRHDCDLREYEEQEKAKLQSGMSANSEVQW 1372
QY 224 -----RLKKDLYSSRSKSLKTLNTE---LDQAKLELSAQKDL--- 257
DB 1373 RTKYETDAIQTELELEAKKLAQRLQDAEESIEAVNSKCSLEKTKORLQGEVDELMD 1432
QY 258 -----QSADQETLSLRKSSDPPGRLPASAATNEVSLVSPAPVE-- 300
DB 1433 VERANSLAANDKFORNDKVLAEKKQYESSQALEGAQKEARSLSTELFKMKNSYEAA 1492
QY 301 -----MMNRLHOPFPFGDEIDNTFTPDVNTPTTSGSOHCPRK-----K 340
DB 1493 LDHLETLKREKKNLQO-----EIS-----DLTEQLGTEGKSHLEKAKKTVESKRFIQ 1542
QY 341 LCLERANSPMONVLKKYKVKSPBSQSLSGQRCVGEIDELAGAFPLFIRNA--VLQCK 398
DB 1543 TALEBAEGTLEHSESKILRYQLLELNQVSEIDRKLAKEDEEM---EQIKRNSQVIDSM 1598
QY 399 QPNTTASRSSTDVARI 416
DB 1599 Q-STLDESEVRSRNDALRV 1615

```

```

RESULT 32
MYH9_CHICK STANDARD; PRT; 1959 AA.
ID MYH9_CHICK
AC P14105;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle (Cellular myosin heavy chain) (NMHC).
GN MYH9.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
CX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestinal epithelium;
RX MEDLINE=90046668; PubMed=2813355;
RA Shober R.V., Conti M.A., Kawamoto S., Preston Y.A., Brill D.A.,
RA Adelstein R.S.;
FT "Cloning of the cDNA encoding the myosin heavy chain of a vertebrate
FT cellular myosin."
FT Proc. Natl. Acad. Sci. U.S.A. 86:7726-7730(1989).
RL -I- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPRINC.
CC -I- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -I- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M26510; AAA48974.1; -
DR PIR: A33977; A33977.
DR HSSP: P10587; IBLI.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Alkylation; Multigene family.
FT DOMAIN 1 788 MYOSIN HEAD-LIKE.
FT DOMAIN 779 808 IQ.
FT DOMAIN 837 1925 COILED COIL (POTENTIAL).
FT NP_BIND 174 181 ATP (POTENTIAL).
FT DOMAIN 654 676 ACTIN-BINDING.
FT MOD_RES 694 694 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
SQ SEQUENCE 1959 AA; 226502 MW; A75C86086FD3A1A1 CRC64;

```

Query Match 7.1%: Score 169; DB 1; Length 1959;
Best Local Similarity 19.1%; Pred. No. 0.27;
Matches 108; Conservative 108; Mismatches 182; Indels 166; Gaps 21;

QY 48 QCRIVGKKTITNLFEDLAOEENV-----LDNF--LNKELDSVKAQLSQK 93
D 1073 ELKIQLSRKE--EEIQALALAVEEBAQKNNMALKIRRELSQITELQDLESERASRKA 1130
QY 94 DREKDSQAIIIDTLRDITLEENNAVESIQNLNKAEMCLSTKK-----QMKFLEQ 144
D 1131 EKQRDLGEELALUKTELDSTAAQDELRSKREOEVTALKTEDEAKTHEAQIOW 1190
QY 145 RO-----DETQAREEAHRLCKMKTMEQ-----IELLOS-----ORSEV 180
D 1191 RQKNSQAIIEELAEQLEQRRKANLEKAKQALESERALSNVAVLLQGDGAHEKRRKV 1250
QY 181 EEMTRDMGV-----GOSAVEOLA-----YYCSLKEYEYENLKEAR 215
D 1251 DAQLOELQVKEFTEGERVKTLEAEHVNLQVELDNVYTGILNDSQSKSIKLADFSALESQ 1310
QY 216 KATGELADRLKDLVSSRSKTK-----TNTLELDAQKLRLRSQKQNLQSDADEITSLR 268
D 1311 QDTQELQLEETRLKLSSTFKQTEDEKNALKEOLEEBAKRNLEKQISVLQOQAVEAR 1370
QY 269 KKSDDPGNLEPASATN-----ETVSRLVFEESPA--PYEMMNPRLHPPFGDEIDLN 318
D 1371 KKMDDGCGCLELAEBAKKLQKDLSELTQRYEKKIAAYDKLEKTRIQQELDIDAVDLD 1430
QY 319 -----TTFDVNTPTPTSGSQH-----CLP 338
D 1431 HQROTASLKKOKKFDLLAEKNISAKVAEERDRAEAEREKETKALSTARALEEATE 1490
QY 339 KKLCLERA-----RSPMNVL-----KKVHKVSKPESQSLSGGRC--VGEIDELAG 384
D 1491 OKAELEKRNKQFRTIEMEDJMSKDDVGSVHELEKAKALQOYEMKTOLELEDELQA 1550
QY 385 APFLIRNAVLGOKOPNRTTAESRSSTDVVRIGFDLGRTKFIOPRDTTIIIPV--PV 441
D 1551 TEDAKLRLEV-----NQAKMAQDFRDL-----GRDQNEKKKQILIRQVRENEV 1596
QY 442 KSKAKSKQVRIKTVSSASQPKLD 465
D 1597 ELEDERKO---RSIYAVARKKLE 1616

RESULT 33
CX05_HUMAN STANDARD; PRT; 1012 AA.
ID CX05_HUMAN
AC 075665; 075666;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein Cxorif (protein 71-7A).
GN CXORF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=98390187; PubMed=9722947;
RA de Concillis L., Marchitello A., Wapenaar M.C., Borsani G.,
RA Gaglio S., Mariani M., Consalez G.G., Zufardi O., Franco B.,
RA Ballabio A., Banfi S.,
RT "Characterization of Cxorif (71-7A), a novel human cDNA mapping to
RT Xp22 and encoding a protein containing coiled-coil alpha-helical
RT domains".
RL Genomics 51:243-250(1998).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y15164; CA75436.1; -;
DR EMBL: Y16355; CA76185.1; -;
DR MIM: 300170; -;
KW Coiled coil; Alternative splicing.
FT DOMAIN 189 557 COILED COIL (POTENTIAL).
FT DOMAIN 622 662 COILED COIL (POTENTIAL).
FT DOMAIN 867 956 COILED COIL (POTENTIAL).
FT VASPLIC 352 367 KYQELMDVYIIRNR -> NFRHLGVCLALGILI (IN
FT VASPLIC 368 1012 ISOFORM 2).
FT VASPLIC 368 1012 MISSING (IN ISOFORM 2).
SQ SEQUENCE 1012 AA; 116670 MW; C2BF4376F89E6738 CRC64;

Query Match 7.0%: Score 168.5; DB 1; Length 1012;
Best Local Similarity 21.2%; Pred. No. 0.14;
Matches 105; Conservative 91; Mismatches 182; Indels 117; Gaps 23;

QY 66 LAQEEENVLDNF-----FLKNELDSVKAQLSQKD-----REKDSQAIIIDTLRDT 110
D 328 LRQDQNKSPFEETRYDRKLKNEL--LKYQLELDKDYIIRTNRLIEDEKKNKNAHLOEE 385
QY 111 LEERNATVESIQNLNKAEMCLSTLKKOMK--LEORODETQAREEAHRLCKMKT-- 166
D 386 LIAINSKKEELNOSVNRV-----KELELEESVKAQSLAITKQNLMLNEKVEKMSDY 437
QY 167 -----EQLELLQSQ--RSEVEEM-----IRDMGVQSAVEQALAVYCVSLK-- 208
D 438 SLKEEKELELAQNKLKQOLEESRNENLRILNLAQAPALAVQKELRKAEAIVAEH 497
QY 209 ENLKEARKATG-----ELADRLKDLVSSRSKTKTNTLELDAQKLRLRSQKQNLQSD 259
D 498 EEFESCRAQLKQLODELIEHSAQLKAQILGKASVSLTTQVADLQLOKQOTYLENEV 557
QY 260 -----ADQETSLRKKSDDPGNLEPASATNETVSRLVFEESPAVE-----MMNRL 306
D 558 YCNPKQSVIDRSVGNLN-----GNVVPK--NGEISGDFLNNFKQENVLAVRASRI 608
QY 307 HQPF-----GPEIDLNTFDVNTPTPTSGSQHC-----LPKKCLERASPMQ-- 351
D 609 TNYPTAWVEGSSPDSDFEAVNTKARVKELOEAEERLEKAFRSYHRRVIRKNSAKSPLA 668
QY 352 -----NVLKKYHKV--SKPESQSLSGGRCVGE-----LDELAGAFPLFIRN 392
D 669 SPSIHLLEAFKKNITSSSPENHI--FGEDRVYSEQOVGTLEERNVYVALTGSASRLRG 727
QY 393 AVLQKOPNRTTAESRSSTDVVRIGFDLGRTKFIOPRDTTIIIPVVKSKAKSKQVR 452
D 728 GTSSRRLSTPLPKAKRSLF--SEMYLEGL--GRSHIASPSPCPDRPLP--SPTESRHLS 783
QY 453 IKTVSSASQPKIDTF 467
D 784 IIPVSSPEQKVGly 798

RESULT 34
MYH2_HUMAN STANDARD; PRT; 1941 AA.
ID MYH2_HUMAN
AC 090KX2; Q16229; Q14522;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, adult 2 (Myosin heavy chain IIA)
DE (MyHC-IIa).
GN MYH2 OR MYHSA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 ON NCBI_TaxID=9606;
 OX [1]
 RX SEQUENCE FROM N.A.
 RP TISSUE-Skeletal muscle;
 RA MEDLINE=99318669; PubMed=10388558;
 RT Weis A., Schiaffino S., Leinwand L.A.;
 RL "Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity."; J. Mol. Biol. 290:61-75(1999).
 RN [2]
 RP SEQUENCE OF 1711-1941 FROM N.A.
 RC TISSUE-Skeletal muscle;
 RA MEDLINE=95109625; PubMed=7545970;
 RT Smerdu V., Karsch-Mizrachi I., Campione M., Leinwand L.,
 RA Schiaffino S.;
 RT "Type IIX myosin heavy chain transcripts are expressed in type IIB fibers of human skeletal muscle."; Am. J. Physiol. 267:C1723-C1728(1994).
 RN [3]
 RP SEQUENCE OF 1823-1941 FROM N.A.
 RC TISSUE-Skeletal muscle;
 RA MEDLINE=95270723; PubMed=7751403;
 RX Enlison S., Sant'ana Pereira J., Sargeant T., Young A., Goldspink G.;
 RT "Characterization of human skeletal muscle fibres according to the myosin heavy chains they express."; J. Muscle Res. Cell Motil. 16:35-43(1995).
 RL [1]
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PFM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY AKTYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF111784; AAD29990.1; -
 DR EMBL: S73840; AAC13916.1; -
 DR EMBL: Z32858; CAA83687.1; -
 DR HSSP: P08799; IMND.
 DR MIM: 160740; -
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IO; 1.
 DR Myosin: muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW MultiGene family.
 KM

FT DOMAIN 1 786 MYOSIN HEAD-LIKE.
 FT 787 816 IO.
 FT DOMAIN 845 1941 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT CONFLICT 1844 1844 K -> R (IN REF. 3).
 SQ SEQUENCE 1941 AA; 223043 MW; 681E866F83AEAB3F CRC64;
 Query Match 7.0%; Score 168.5; DB 1; Length 1941;
 Best Local Similarity 19.5%; Pred. No. 0.28; Indels 37; Gaps 7;
 Matches 73; Conservative 81; Mismatches 183;
 QY 55 KTTINKLFFDLAOFEEVNLDAEFLKNELDYSKAKQLSOKDEKRSQAIIIDLRPTLEER 114
 DB 1272 GORLNDL---TAQKGRQTESGERSRDLDEKALVSLSNGKAFVQIIEELKQLEEE 1328
 QY 115 NATVESLONALNKAEMLCSTLKQKFLERQODETKQAREAHRLKCKMKMTWEQTELLQ 174
 DB 1329 IKAKNALAHALQSSRHDDLRREOYEEDQSKAEIQRALSKANTEVAQWRTKYETDAIOR 1388
 QY 175 SORSEVEE---MTRDMGVGSAVEQGLAVYCVSLKK-----EVENLKEARKATGELAD 223
 DB 1389 TE--ELEPAKKKLAQRLQAAEBHVAANAKASLEKTKQRLONEVEDMLDVERTNAACA 1446
 QY 224 RLKQDLVSSRSKLTLTNLTDLQAKLELSAQKDLQADQETLSLKKSDPPGNLEPASA 283
 DB 1447 ALDKQRNFQDILAEWKRCQCEHAELEASQKARSLSLETFLFKINAYEESLDQLETIKR 1506
 QY 284 TNEVTSRLVESPAPVEKMPRLHOPPGDEIDLTTFDVNTPPTQISGSHCLPKIKLC 343
 DB 1507 ENKNLQOEISDLTEQIAGCKRIHE-----LEKIKQVEQKCELQAL 1550
 QY 344 ERASPMQNVLKVKYKVPESQSLSGQRCVGELEDEELAGAFPIFINAV--LQKKQNNR 402
 DB 1551 EBAESLSHEGKTLRIQLELNQKSEVDKRIAEKDEI---DQLKRNHRIVESMOST 1606
 QY 403 TTAESRSSTDVARI 416
 DB 1607 LDAEIRSRNDIAIRL 1620
 RESULT 35
 TPR_HUMAN STANDARD; PRT; 2349 AA.
 ID TPR_HUMAN
 AC P12270;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nucleoprotein TPR.
 GN TPR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 OX [1]
 RX SEQUENCE FROM N.A.
 RP MEDLINE=93064711; PubMed=1437155;
 RA Mitchell P.J., Cooper C.S.;
 RT "The human tpr gene encodes a protein of 2094 amino acids that has extensive coiled-coil regions and an acidic C-terminal domain."; Oncogene 7:2329-2333(1992).
 RL [2]
 RP REVISIONS AND CHARACTERIZATION.
 RX MEDLINE=95096166; PubMed=7798308;
 RA Byrd D.A., Sweet D.J., Pante N., Konstantinov K.N., Guan T., Saphire A.C.S., Mitchell P.J., Cooper C.S., Abi U., Gerace L.;
 RT "tpr, a large coiled coil protein whose amino terminus is involved in activation of oncogenic kinases, is localized to the cytoplasmic RT surface of the nuclear pore complex."; J. Cell Biol. 127:1515-1526(1994).
 RL [3]
 RP SEQUENCE OF 1-142 FROM N.A.
 RX MEDLINE=88262257; PubMed=3387099;

RA King H.W.S., Tempest P.R., Merrifield K.R., Rance A.J.;
RT "cpr homologues activate met and raf.";
RL Oncogene 2:617-619(1988).
CC -1- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILIBRILS OF THE NUCLEAR PORE
CC COMPLEX IMPLICATED IN NUCLEAR PROTEIN IMPORT. ITS AMINO TERMINUS
CC IS INVOLVED IN ACTIVATION OF ONCOGENIC KINASIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF THE NUCLEAR PORE
CC COMPLEX. THE ASSEMBLY OF THE NPC IS A STEWIE PROCESS IN WHICH
CC TRP-CONTAINING PERIPHERAL STRUCTURES ASSEMBLE AFTER OTHER
CC COMPONENTS, INCLUDING P62.
CC -1- TISSUE SPECIFICITY: HIGHEST IN TESTIS, LUNG, THYMUS, SPLEEN AND
CC BRAIN, LOWER LEVELS IN HEART, LIVER, AND KIDNEY.
CC -1- DISEASE: INVOLVED IN TUMORIGENIC REARRANGEMENTS WITH THE MET, TRK
CC OR RAF GENES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66397; CAA47021.1; -;
DR EMBL; Y00672; CAA6681.1; -;
DR PIR; S00928; S00928.
DR MIM; 189940; -;
KW Colled coil; Proto-oncogene; Chromosomal translocation;
KW Nuclear protein; Transport.
FT DOMAIN 78 360 COILED COIL (POTENTIAL).
FT DOMAIN 422 571 COILED COIL (POTENTIAL).
FT DOMAIN 575 628 COILED COIL (POTENTIAL).
FT DOMAIN 758 805 COILED COIL (POTENTIAL).
FT DOMAIN 834 869 COILED COIL (POTENTIAL).
FT DOMAIN 934 979 COILED COIL (POTENTIAL).
FT DOMAIN 1004 1064 COILED COIL (POTENTIAL).
FT DOMAIN 1138 1166 COILED COIL (POTENTIAL).
FT DOMAIN 1196 1241 COILED COIL (POTENTIAL).
FT DOMAIN 1262 1304 COILED COIL (POTENTIAL).
FT DOMAIN 1354 1434 COILED COIL (POTENTIAL).
FT DOMAIN 1476 1595 COILED COIL (POTENTIAL).
FT DOMAIN 1527 530 POLY-SER.
FT DOMAIN 1833 1836 POLY-GLU.
FT DOMAIN 1957 1964 POLY-ASP.
FT DOMAIN 2295 2298 POLY-SER.
SQ SEQUENCE 2349 AA; 265600 MW; AFDD685CEDCA9EF CRC64;

Query Match 7.0%; Score 168.5; DB 1; Length 2349;
Best Local Similarity 20.5%; Pred. No. 0.34;
Matches 105; Conservative 68; Mismatches 169; Indels 169; Gaps 19;

QY 50 RIQVGKTT---INKLFDLAOEENVIDA-----EFLKNEIDSV---KAQLSQ 92
DB 1405 KVRTEKETIQKDLDAKIIDIEKVTIQVKRIGRYKTOYEELKAQODKWTETSAQSSG 1464
QY 93 KDRERDSQAIIDRLDRLLEENATVESLQNALNKAEMLCSTIKKQMFLOQRQDETQOA 152
DB 1465 DHQGHVSVOEKQELKETLNOAETKSKSL---SOVENLOKTLKSTETEARNLQOTVOL 1521
QY 153 REBAHRLCKMKTMEQIELLQSOSEYEEMIRDMGV--GQSAVEQLAVYCVSLKKEVENL 211
DB 1522 QSELSRLQODQDRTOGEOLRQOTTEKEKTKRAIVAKSKIALHVAVKNQOLTENNEL 1581
QY 212 KEARKATLEADRLKLDVSRSKLTINTELDQAKLELSAQ--KDLQASDQ----- 262
DB 1582 KQGNALDQOQDELDRITALKSOYEGRISRLER---ELREHQRHLRQDEPOEPSNKV 1638
QY 263 -----EIT-----SLRKSDDPPGNLE----- 279
DB 1639 PEQQRQITLTKTTPASGEGINASTSDPTANTIKPTPVVSTPSKVTAAAMAGNKSTPRASIR 1698
QY 280 ----PASATNNTV-----SRLVESPAPVEMMNPRLHQPPFGD----- 313

DB 1699 PMVTPATVPTTTPATVPTTQVQSEAMQSEGPVE-----HYVFSGTSVSRSTS 1752
QY 314 -----EDLNTTDDVNTTPPTQTSQCHLCKLCLERASPM-QNVLKYNK- 359
DB 1753 PMVQPSISQPLTAVQQTQATAFVQPTQOSHPO-----IEPANOLSNIVEVOSS 1804
QY 360 -VSKPEQSLSIGGRCVGEIDELAGAFPLFIRNAVLGOKQNPRTAESRSST-----DVV 414
DB 1805 PVPEPSTSTAVFG-----TVSATSTSSSLPKRTREEDSTTEAD-- 1844
QY 415 RIGFDGLGGRKTFQPRDPTIIRPVVSKA 445
DB 1845 -----QVSDDTFVEMPLPKLKS 1861

RESULT 36
MEPL_LYCES
ID MEPL_LYCES STANDARD; PRT; 697 AA.
AC P93203;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MAR binding filament-like protein 1.
GN MEPL.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
OC Asteridae: euasterids I: Solanales: Solanaceae: Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV_VENT CHERRY; TISSUE=fruit;
RX MEDLINE=97112038; PubMed=8953774;
RA Meier I., Phelan T., Gruissem W., Spiker S., Schneider D.;
RT "MEPL, a novel plant filament-like protein with affinity for matrix
RT attachment region DNA.";
RL Plant Cell 8:2105-2115(1996).
CC -1- FUNCTION: BINDS DNA. INTERACTS WITH CHROMATIN VIA MATRIX
CC ATTACHMENT REGIONS (MARS). LIKELY TO PARTICIPATE IN NUCLEAR
CC ARCHITECTURE BY CONNECTING CHROMATIN WITH THE NUCLEAR MATRIX AND
CC POTENTIALLY WITH THE NUCLEAR ENVELOPE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y07861; CAA69181.1; -;
KW Nuclear protein; DNA-binding; Coiled coil.
FT DOMAIN 125 681 COILED COIL (POTENTIAL).
FT DOMAIN 10 15 POLY-SER.
SQ SEQUENCE 697 AA; 79516 MW; 700A56D8D6A7E49 CRC64;

Query Match 7.0%; Score 168; DB 1; Length 697;
Best Local Similarity 21.0%; Pred. No. 0.098;
Matches 74; Conservative 66; Mismatches 132; Indels 80; Gaps 10;

QY 62 LFPDLAEEEN---VLDAEF-----LKNELDSVAQAQSKQKREKDSQAIIIDRLILE 113
DB 394 LIADLTREKENLRRVWVADLQVNVKLEQIEVTQDESNSRSEVSDITVQLEQRLSSK 453
QY 114 RNATVESLQNALNKAEMLCSTIKKQMFLOQRQDETQOAREBAHRLCKMKTMEQIELL 173
DB 454 LEREYSKIQ-----MELEETRASLQNRNIDETKHSSE-----LL 486
QY 174 QSORSEVEEMIRDMGVGQSAVEQLAVYCVSLKKEVENLKEARKATGELADRLKLDVSSR 233

CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB012247; BAB02666.1; -
 KW Nuclein protein; DNA-binding; Coiled coil.
 FT DOMAIN 140 692 COILED COIL (POTENTIAL).
 FT DOMAIN 22 35 POLY-SER.
 SQ SEQUENCE 727 AA; 82071 MW; 5F4F9A0F7F623CAA CRC64;
 Query Match 7.0%; Score 167.5; DB 1; Length 727;
 Best Local Similarity 20.6%; Pred. No. 0.11;
 Matches 92; Conservative 80; Mismatches 155; Indels 119; Gaps 15;
 Oy 68 QEEENVLADEFK-----NELDSVKAQLSOKREKDSQAIIIDLTLEERNATVES 120
 Db 187 EOEERKKKEVEKAKEBOLSLINOLNSAKDLVELGRELSSEKKLCKEKLD-----QIES 239
 Oy 121 LONALNKA-----EMLCSTLKKOMKFLEROD-----ETKQAREAHRLCKM----- 163
 Db 240 LENSLSKAGEKALFETLREKLDVEGLDRINLSLELKDSEKQAFNALSIAKKEAE 299
 Oy 164 -----KTMQEJELLQSORSEVEEMIRDMGVGSAVEQLAVYCVSLKKEYENL 211
 Db 300 LKELNSIYTOTSRDLAEKLEIKOQKEELRTQSELDSKNSAIEBLRITTLVAEKESY 359
 Oy 212 KE-----ARKANGELADRLKDLVSSR-SLTKLTNLNLDQAKELSLAQKDLSA 260
 Db 360 IQKLDLSIKDYSAKLVISETQAAADAEILSRKEDTQQLNENLDL-----ALDDVAKS 412
 Oy 261 DOIETSLRKKSDPPGNLEPASATNETVSRVLFESPAPVENMPRLHQPFGEDIDLNTT 320
 Db 413 KDKVADLLEKEDSKRMLDIETTYKNLRHLEGGTKTLQASRDVVS-----DLFTM 464
 Oy 321 PDVNTPPQTSGSOHCLPKKLCLEPARSPMONVNLKRVKSKPESQSLSGQRCVGELE 380
 Db 465 LDES-----RALC-----SKLSESLAI-----VHE 484
 Oy 381 ELAAGFPIFIRNAVVGOKOPRTA-ESRSSDVVRIGFDGCGTKRIQPRDTIIRPV 439
 Db 485 EMKEAKERYERN-LDAERQKNEISASLELKKDLRRRVKDELEGVTHELKSS----- 536
 Oy 440 PVKSKAKSKOKVRIKTVSSASQPKLD 465
 Db 537 -VKNGSLQKLELVEIKKVEYSKKELE 561
 RESULT 39
 MYSB_CAEEL
 ID MYSB_CAEEL STANDARD: PRT; 1966 AA.
 AC P02566;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain B (MHC B).
 GN UNC-54 OR MYO-4.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83273600; PubMed=6576334;
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the Caenorhabditis elegans unc-54
 RT myosin heavy chain gene are not separated by introns.";

RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
 RN [2]
 RP SEQUENCE OF 850-1966 FROM N.A.
 RX MEDLINE=82272395; PubMed=720124;
 RA McLachlan A.D., Karn J.;
 RT "Periodic charge distributions in the myosin rod amino acid sequence
 RT match cross-bridge spacings in muscle.";
 RL Nature 299:226-231(1982).
 RN [3]
 RP SEQUENCE OF 1876-1966 FROM N.A.
 RX MEDLINE=83232892; PubMed=6571695;
 RA Wills N., Gesteland R.F., Karn J., Bolten S.,
 RA Waterston R.H.;
 RT "The genes sup-7 X and sup-5 III of C. elegans suppress amber
 RT nonsense mutations via altered transfer RNA.";
 RL Cell 33:575-583(1983)
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
 CC C.ELEGANS.
 CC -1- WALL: MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J01050; AAA28124.1; -
 DR EMBL: V01494; CAA24738.1; -
 DR PIR: A02992; MMRK.
 DR HSSP: P08799; IMND.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; Myosin_head.
 DR Pfam: PF00063; myosin_head.1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head.1.
 DR SMART: SM00242; MySc; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 850
 FT MYOSIN HEAD-LIKE.
 FT DOMAIN 851 1966
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1165 1176
 FT ALPHA-HELICAL TAILPIECE (S2).
 FT HINGE.
 FT DOMAIN 1165 1966
 FT LIGHT MEROMYOSIN (LM).
 FT NP_BIND 177 184
 FT ATP (BY SIMILARITY).
 FT ACTIN-BINDING.
 FT ACTIN-BINDING.
 FT METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 128 128
 FT ACTIN-BINDING.
 FT MOD_RES 705 705
 FT ALKYLATION (SH-1).
 FT MOD_RES 715 715
 FT ALKYLATION (SH-2).
 FT CONFLICT 1337 1337
 FT E -> R (IN REF. 2).
 FT CONFLICT 1880 1880
 FT I -> L (IN REF. 2).
 SQ SEQUENCE 1966 AA; 225125 MW; B66F0BB2FE27B67F CRC64;

Query Match 7.0% Score 167.5; DB 1; Length 1966;
 Best Local Similarity 19.9%; Pred. No. 0.32; Indels 145; Gaps 17;
 Matches 85; Conservative 71; Mismatches 126;

55 KKTITNKLFFDIAQEEVNLDAEFLKNEIDSVKAOISQKREKRSQAIIIDTLRPLEER 114
 DB 1011 GEINRKLMEDEQSEEDK-----GNHONKRYAKLEQ-----TLDLESLERE 1053
 OY 115 NATVSLONALKAKEMLCSTLKQKPFLEQRODER-KQAREPAHRLKCKMTMEDIEILL 173
 DB 1054 KRARADLKQKRRVE-----GEIKIQENIDESGRHNDENNLLKKKESLSVSSRL 1106
 OY 174 GORSEVEEMIDMGVSGAVBOLAVCVSLKKEVENLKEARKATGELADRLKDLVSRSR 233
 DB 1107 EBEOLVSLKQKIDGOSRISE-----LEELENERQSSK-----ADRAKSLP---Q 1152
 OY 234 SKLTITNLELDQ-----AKLELSAOKDLQSA---DOETISLRKSSDDPP 275
 DB 1153 RELEELGEKIDEGGATGAQAVEVNNKREAEELAKLRDLEAMNHNENGLGKKHTDAY 1212
 OY 276 GNLBPASATNETVSRVLESPRPVEMANPRLLHOPFGDEIDLTFTDVTPTQTSQSH 335
 DB 1213 AELT-----DQLDQ----- 1221
 OY 336 CLPKKICLERARSPQNVLLKVKYKSPESOLSLGQRCVGLDEBLGAPPLFIRNAV 395
 DB 1222 -----LNKAKA-----KVEKDKAQAVRAEDLAOLDDETSGK-----LNNEKL 1260
 OY 396 GQ-----KOPNRTTAESSSTVDVVRIGFDLGGRTGFIPRODTTITRVPVSKASK 448
 DB 1261 AKQFELQLELQSKADEQSRQLQD-----FTSLKGR---LHSENGDLVROLE-DAESQVN 1311
 OY 449 QKVRITK 455
 DB 1312 QLTRLS 1318

RESULT 40
 MYHB_CHICK STANDARD: PRT; 1978 AA.
 AC P10587;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, gizzard smooth muscle.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88118918; PubMed=2892941;
 RA Yanagisawa M., Hamada Y., Katsuragawa Y., Imamura M., Mikawa T.,
 RA Masaki T.;
 RT "Complete primary structure of vertebrate smooth muscle myosin heavy
 RT chain deduced from its complementary DNA sequence. Implications on
 RT topography and function of myosin.";
 RL J. Mol. Biol. 198:143-157(1987).
 RP [2]
 RP REVISIONS.
 RA Masaki T.;
 RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE OF 1-203.
 RX MEDLINE=88032919; PubMed=3312184;
 RA Matsu G., Onishi H., Yajima E., Matsuda G.;
 RT "Amino acid sequence of the amino-terminal 24 kDa fragment of the
 RT heavy chain of chicken gizzard myosin.";
 RL J. Biochem. 102:133-145(1987).
 RN [4]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 1-818, AND REVISIONS 204-215.
 RX MEDLINE=98412652; PubMed=9741621.
 RA Dominguez R., Freyzon Y., Trybus K.M., Cohen C.;
 RT "Crystal structure of a vertebrate smooth muscle myosin motor domain
 RT and its complex with the essential light chain: visualization of the
 RT pre-power stroke state.";
 RL Cell 94:559-571(1998).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PFM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X06546; CAA29793.1; -.
 DR PIR: S03166; S03166.
 DR PDB: 1BR1; 09-SEP-98.
 DR PDB: 1BR2; 09-SEP-98.
 DR PDB: 1BR4; 09-SEP-98.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Spectrin.
 DR InterPro: IPR002017; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00663; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PSS0096; IQ; 1.
 KW Myosin: muscle protein; coiled coil; thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;
 KW Multigene family; 3D-structure.
 FT INIT_MET 0
 FT DOMAIN 1 790 MYOSIN HEAD-LIKE.
 FT DOMAIN 791 820 IQ.
 FT DOMAIN 849 1978 RODLIKE TAIL (S2 AND LM DOMAINS).
 FT DOMAIN 849 1978 COILED COIL (POTENTIAL).
 FT NP_BIND 176 183 ATP.
 FT DOMAIN 666 688 ACTIN-BINDING.
 FT DOMAIN 767 781 ACTIN-BINDING.
 FT MOD_RES 1 1 BLOCKED.
 FT MOD_RES 127 127 METHYLATION (TR1-) (POTENTIAL).
 FT MOD_RES 706 706 ALKYLATION (SH-1).
 FT MOD_RES 716 716 ALKYLATION (SH-2).
 FT CONFLICT 127 127 MISSING (IN REF. 3).
 FT CONFLICT 204 215 KDTSTQGPSFS -> RTPASLKVHLFP (IN REF. 1).
 SQ SEQUENCE 1978 AA; 228663 MW; B7B6C92385273D93 CMC64.

Query Match 7.0% Score 167.5; DB 1; Length 1978;
 Best Local Similarity 23.0%; Pred. No. 0.32;

THIS PAGE BLANK (USPTO)